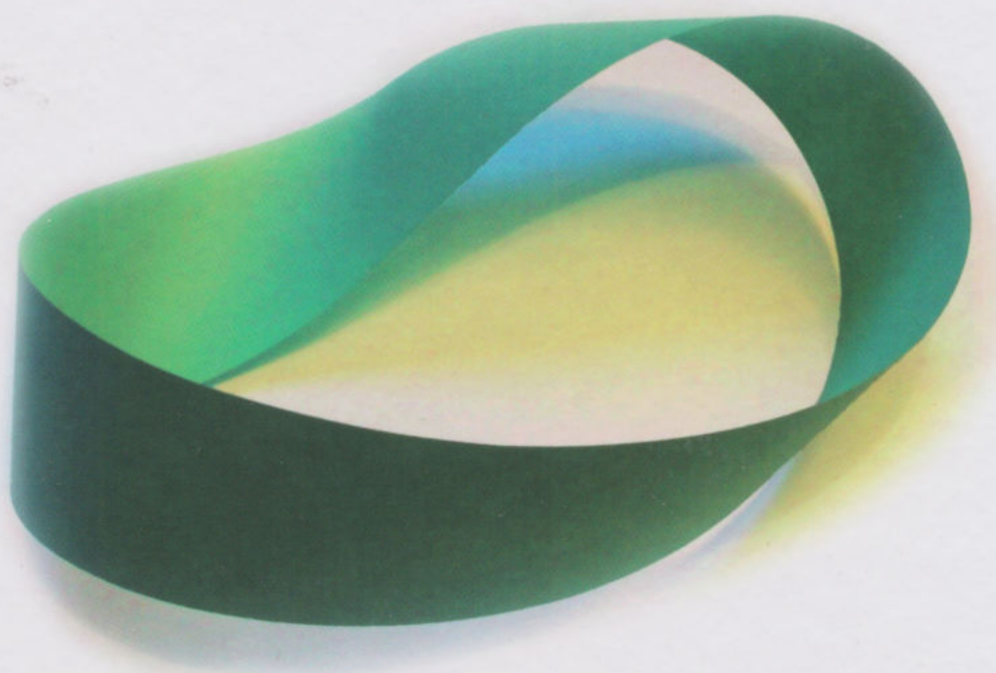
Into Infinity

Discovering a world without end



Everything is mathematical

Into Infinity

Into Infinity

Discovering a world without end

Enrique Gracián

Everything is mathematical

Published by RBA Coleccionables, S.A.
c/o Hothouse Developments Ltd
91 Brick Lane, London, E1 6QL

Localisation: Windmill Books Ltd.
Photograph credits: iStockphoto

ISSN: 2050-649X

Printed in Spain

Contents

Preface

The French writer Alphonse Allais (1854–1905) claimed that "Infinity is long, above all towards the end," which, as well as being quite funny, expresses that our vision of infinity is always associated with a certain proximity. Put another way: We can only see it from here, from a place where, as finite beings, we find ourselves within the limits of our perceptions. When we look into the distance, we begin to lose track of ourselves and turn instead to philosophical ruminations and conjecture. Usually, we opt for an intellectual answer, as opposed to a simple acceptance of the impossibility of the issue. Hence there is nothing strange about the fact that the concept of infinity has had, and continues to have, a broad appeal that interests philosophers, theologians, scientists and, of course, mathematicians alike. Between these four fields of human endeavour, the boundaries are not always clear...

For most people, their first reaction to the idea of infinity is a feeling of vertigo, of being faced with something that, as hard as we try, will invariably escape us. Perhaps this is one of its greatest attractions – infinity makes anything possible. Nevertheless, the history of mathematics contains a chapter on the 'mathematics of infinity'. This amorphous, challenging subject has been converted into just another mathematical object, just another number or geometrical form.

A mathematical object is really nothing more than a good definition. In this respect, a mathematician is a hunter: he or she explores unknown territory, stalking and observing prey, waiting until the target is clear before taking aim.

And this is the story of infinity in mathematics, although this 'prey' took more than 3,000 years to quell. Infinity hunters carved a path between dogma and paradox, ascended the highlands of Greek philosophy, crossed the green fields of religion and crept around the dark forest of secretive sects. At the same time, the subject involved a journey into geometric shapes and through the labyrinth of numbers. The quest required searchers with a very special skill set.

As such, we can follow the path to infinity from inside the minds of the best thinkers from all cultures, be they philosophers, theologians, doctors, mathematicians, or even madmen, in an adventure that is not without its dangers. Some travellers have paid the price with mental illness, others literally laid their lives on the line and ended up being punished by extreme cults or burned at the stake by intransigent religious authorities. All this for something that is nothing more than an idea. However, we know that even an idea can change the way we see the future, contemplate beyond our own existence – and, consequently, dislodge the foundations upon which the beliefs of all cultures rest.

The fact is that infinity and our understanding of it informs our world view, and as such it is of interest not only to mathematicians but also to philosophers. But we must not forget that both points of view can live in harmony, or as the French mathematician Jean Charles de Borda (1733–1799) put it: "Without mathematics, we do not get to the bottom of philosophy; without philosophy, we do not get to the bottom of mathematics; without both, we do not get to the bottom of anything."

Chapter 1

What is Infinity?

The concept of infinity is inherent to human thought. It is likely we are born with a vague mental concept of the infinite, which we quickly come to associate with its opposite, the clear perception of the finite in our surroundings. In philosophy and theology, reflecting on the infinite is purely circumstantial – you can take it or leave it. However, in mathematics, the investigation of infinity will always be necessary.

Infinity in everyday life

There is a famous legend about a mathematics teacher who had to explain infinity to his class for the first time. He took a box of chalk, removed a stick and began to draw a line on the blackboard. Upon reaching the edge, he continued drawing the line on the wall, all the way down to the floor, and kept going. Still drawing, he went out through the classroom door and disappeared at the end of the corridor, still holding the chalk. The pupils were amazed and sat waiting for something to happen. Shortly after, the bell rang to announce the end of the class. The teacher had disappeared. The last person to see him was a porter who watched him make his way down the street leaving a white line along the walls of the houses in his wake. After three days, the management of the school decided to call in a replacement. Some months later, the teacher suddenly appeared halfway through the mathematics class. He had a rucksack, a long beard and, of course, a small piece of chalk in his hand. He strode back into the room, drawing a line on the floor, then up the wall, until reaching the blackboard, where he stopped. Clearly exhausted, he addressed his students and said: "This is an incredibly long line, but it is nothing compared to infinity."

It is not known whether the school decided that the teacher should continue his lesson or whether they made an appointment with the men in white coats. Nor is it known if his students had any better understanding of what infinity meant despite his great efforts. What we do know, however, is that infinity does demand an exceptional – although not necessarily so extreme – treatment.

There are other increasingly fantastical stories that try to help us perceive of infinity. In religious contexts, it was common to refer to the infinity of time, more commonly known as eternity, when speaking of divine punishment. Purgatory may have lasted for a long time, but it was not eternal. Punishment in Hell, however, lasted for an infinite period of time. To give a rough idea of what this meant, gargantuan tasks were proposed, such as collecting all the grains of sand from immense beaches, picking up them up at a rate of one every hundred years.

One of the most curious examples of eternity was the following: Imagine the Earth was a compact ball of steel, and a dove brushed against its surface once every million years. When the sphere had been worn down to a tiny point, eternity would have passed (although it would be more correct to say 'an' eternity). All these examples were explained to children to give them an idea, albeit an unfortunately menacing one, of the immense magnitude of infinity.

Illustration by Gustave Doré for Inferno, *the first part of Dante's* Divine Comedy. *Going to Hell meant eternal suffering, an infinite punishment.*

I had my first 'vision' of infinity as a child, the first time I found myself caught between two parallel mirrors inside a lift. "What's this?" I asked. My father took my hand and replied: "That's infinity." From then on, for me infinity was a landscape, both frightening or fascinating depending on how I looked at it. Even now it is best to travel through it holding hands.

For all of us, infinity is a completely inaccessible place, which at best invokes a certain apprehension and at worst creates a cosmic horror. However, on the other hand, the alternative to infinity is not particularly encouraging either. If the universe is finite, what lies beyond its bounds? The answer: Nothing. With a capital N. A concept perhaps more disturbing than infinity.

A dictionary definition

The word 'infinity' is frequently used in everyday language: 'infinite space', 'an infinite number of times', 'an infinite period of time', 'to have infinite patience' are all common expressions. We all understand what they mean, provided we do not give the matter too much thought. If we were to, it would soon become clear that our ability to comprehend infinity becomes exhausted straight away; we are only able to imagine 'snapshots' of the reality, which are of little or no help. And the fact is that such a concept is unavoidably philosophical: to think of infinity is to think of philosophy, a subject for which one must have a certain predisposition and, above all, a starting point. Under these circumstances, the easiest course of action is to turn to a dictionary:

infinity

(From Latin *infinītus*).

1. adj. That does not have or cannot have an end or finish.

2. adj. Of a very great number.

3. noun. A place that is imprecise in terms of distance and vagueness. *The street vanished into infinity.*

4. noun. In photography, the furthest range at which an object is effectively in focus when the lens is set for the greatest possible distance.

5. noun. *math.* Value that exceeds any assignable quantity.

6. noun. *math.* The symbol (∞) used to express this value.

7. adv., noun. Excessively, a lot.

Let's now consider these definitions in more detail, without seeking to challenge their interpretations, but with the aim of getting as close as possible to the mathematical meanings. The first explains that infinity refers to that which does not have an end. This statement does not say what it does not have, but what it might not have. Admittedly this is a subtle difference, since by stating infinity does not have an end, we are making two claims, one implicit, which states that infinity exists, and another, which is explicit, that states it does not have an end. The second statement, by saying infinity cannot have an end, affirms that in the event it does exist it cannot be finished. Perhaps this difference may seem like splitting hairs, but it involves a crucial concept – the difference between potential infinity and actual infinity. We shall have to deal with this later.

The second, third and last definitions are qualitative ones that make reference to perception or feelings. The fourth, which deals with a technical matter, comes from a geometrical context that represents an important milestone in the history of mathematics, whereby infinity was interpreted as a point at which two parallel lines meet. The sixth definition makes reference to the sign used for symbolising infinity, which we shall also discuss further on. The fifth, 'a value that exceeds any assignable quantity', is the one closest to the mathematical concept of infinity.

The basic concepts we have observed so far are related to the idea of infinite progress, both in terms of space and time, but the object to which they are applied is too vague, and the term 'infinity' is used more as an adjective than a noun. When we talk of 'eternal love,' we are referring to the concept of infinity in terms of time, whereas in fact, what we really mean is extreme fidelity. On the other hand, if we say that the universe is infinite, we are referring to 'immensity in terms of space'. Things continue to be somewhat vague. One example leads us to contemplate the sky on a moonless night, with millions of stars and a supposed deep black background, which we give the unsettling dimension of infinity. It is clear that the first thing we must do if we wish to tackle the matter of infinity, is set our sights on a concrete object. Although it may seem paradoxical because of the supposedly abstract nature of mathematics, we find the best starting point in the natural series of numbers or, more correctly, the series of natural numbers.

There is nothing more natural than a natural number, and all advanced cultures throughout the world understand what we mean when we talk of the series of natural numbers (1, 2, 3,...). But when does the series end? The immediate reply is never. But why? Because we can always add another number. The answer is correct, so correct that, as we shall see later, it encompasses an extremely precise definition

THE SYMBOL FOR INFINITY

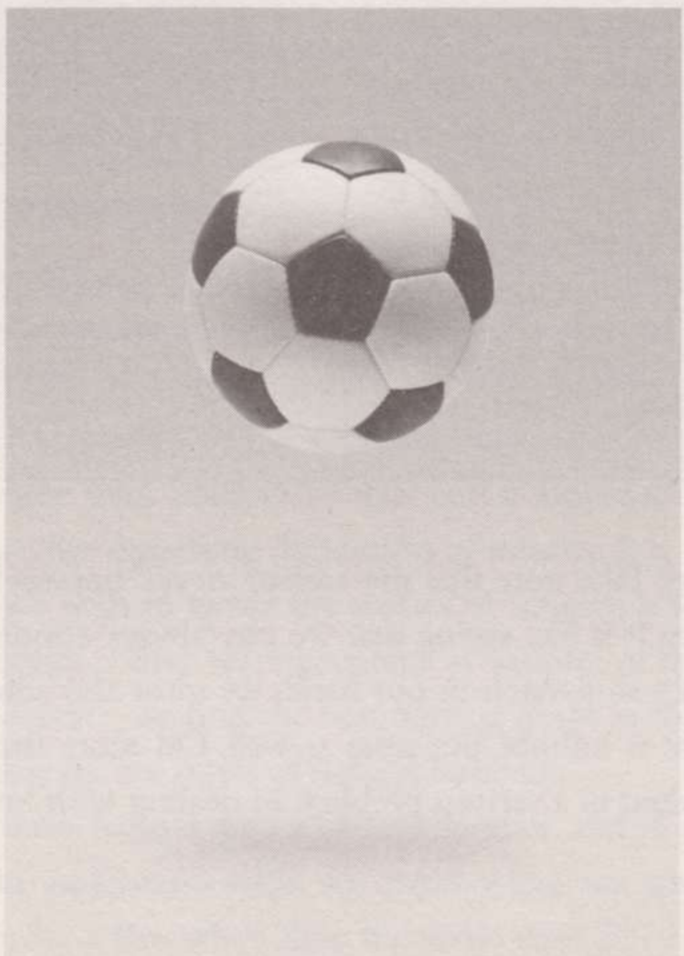
The circle that hangs over the images of saints represents eternity. In Latin, the word *caelum* means both 'heaven' and 'circle', and is, as an endless curve, a path that may be travelled an infinite number of times during an infinite period of time, and hence a possible representation of eternity. In the same way, in certain pagan contexts, the symbol for infinity has been used as a sign of sanctity, replacing the circle. In most packs of tarot cards, the symbol for infinity can be found on the first card, above the head of the Magician. The symbol, which many people falsely interpret as a 'figure of eight on its side', corresponds to a curve named the 'Lemniscate of Bernoulli'. It was introduced by the British mathematician John Wallis (1616–1703). Another suggestion is that the sign came from the italicised form of the letter M (the symbol for a thousand), used by Wallis, who was a linguist, to represent an extremely large number.

In the tarot card for the Magician, the symbol for infinity hangs above his head.

of the term infinity. Take note that the answer 'never' implies both a temporal and numerical property. It is like saying that we can 'always' continue adding numbers. If we do this with a stopwatch in our hand, we must also admit it is not only the numeric series that is infinite, but time as well. On more than one occasion, this confusion has resulted in a serious problem in dealing with infinity.

Extremely large and extremely small

Consider this thought experiment. Imagine we have a ball with the following property: Every time it bounces on the ground, it rises to half the height from which it fell. Hence, if we let it fall from two metres, it will rise to one metre on the following bounce, and then 50 cm, and so on. Let us now consider the following problem. The ball is dropped from 10 metres. When it has come to a stop (i.e. when it stops bouncing) how far will it have travelled? We can't say that this problem is beyond reality. It is literally child's play to check that at some point the ball will stop bouncing. It cannot go on making small bounces forever. On the other hand, however, we can interpret the path taken by the ball as being infinite, since the potential for dividing its height in half is endless, giving us an increasingly smaller height after each step, all the way down to as small a number as we can imagine.



Will a ball stop bouncing at some point in time or will it continue forever, with increasingly smaller and less perceptible bounces?

This is a typical paradox associated with infinity, which we shall consider in greater depth further on. For now though, it suffices to introduce us to something new: the infinitely small.

As such, the idea of immensity is not the only concept we can have of infinity; there is also the infinitely small. Consider a line segment divided into two parts. One of them can be divided into another two, and so on, without stopping. At least in theory, we can make an infinite number of divisions, giving us increasingly smaller segments. Does this process come to an end? No, since just as we have noted, it is always possible to add another number to the series of natural numbers so it is also always possible to divide a segment (or number) in half. As such, infinity refers both to the extremely large, and the extremely small. These two actions are referred to as 'infinite extension' and 'infinite divisibility'.

Apeiron

The first attempt to tackle the subject of infinity can be found, as always, by going back to the foundations of philosophy in classical Greek culture. As is well known, one of the many virtues of the Greek philosophers was the creation of their own philosophical language. They created a specific word to represent an idea, thus establishing what we would now call philosophical terminology. This had equal or greater precision than scientific terminology, which is ultimately derived from it. For our purposes, the key word was *apeiron*, derived from *perata*, which means the 'limit of something' and hence, what does not have a *perata* is *apeiron*, infinite, limitless.

In the context of Greek philosophy, this limitlessness takes on a special meaning that refers not just to the idea of infinite extension, as we understand it in everyday use, but to the origin of all that exists. The underlying idea is that anything that exists does so as a function of its limits. If we think of any object, such as a table, the first thing we observe, even before its functionality, is the limits that define it and separate it from what surrounds it. This idea applies to both inanimate and living entities. What makes a living cell exist is a membrane that imposes limits with the medium that surrounds it. This allows us to say that everything that 'is' exists within its limits and does so thanks to these limits.

As such, the *apeiron* is like an undefinable magma in which everything exists thanks to the appearance of precise limits within it. Hence, being *apeiron* is more defined by the presence of the undefined than the unlimited. Consequently, it is not strange that in addition to a creative power, *apeiron* is also attributed with the power of providing

certain functions and even virtues to things that are created. Hence the idea of *apeiron*, and, consequently, the idea of infinity, has been associated with God from the very beginning and in all kinds of different religions.

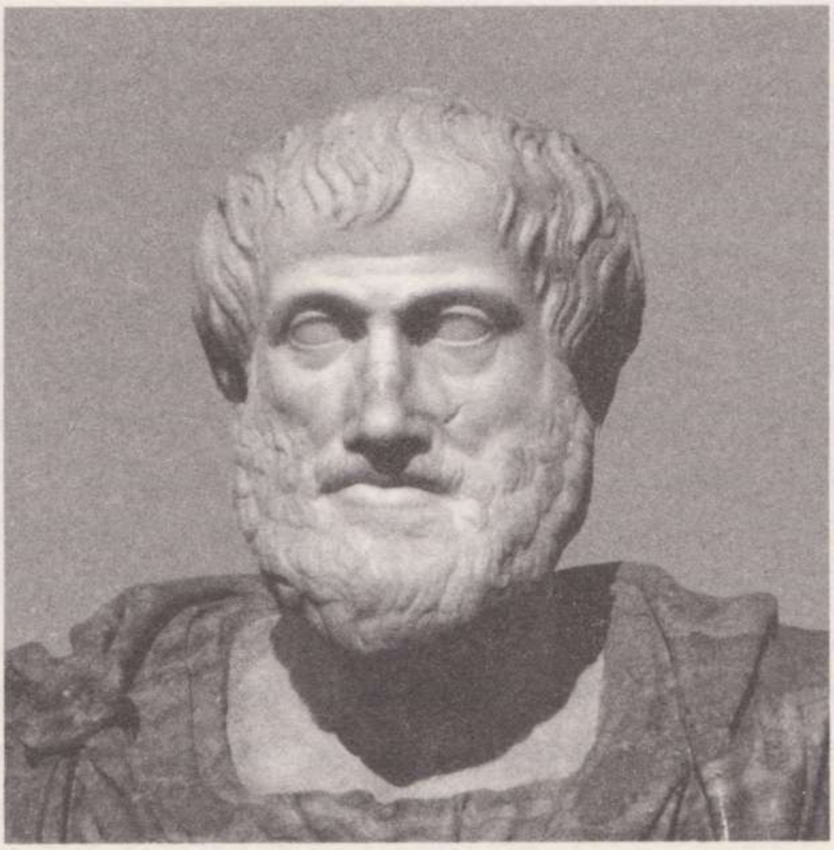
And hence there is a certain ambivalence or contradiction in the term, given that *apeiron*, as the origin of things, has connotations of the original *kaos* and, as such, is associated with evil, an undesired element that does not form part of our existence. This is the origin of our conflicted attitude toward infinity, which can be associated both with unreachable divinity and the disorganised and chaotic forces of evil. The negative aspect of infinity has been perpetuated throughout history. J.L. Borges refers to it thus: "One concept corrupts and confuses the others. I am not speaking of evil, whose limited sphere is ethics; I am speaking of the infinite."

Another use of the concept of *apeiron*, although this time closer to what we commonly understand as infinity, leads us to the idea of a Euclidean space, in the sense of a geometric space without limits. In this respect, and in keeping with the thoughts of Plato, Aristotle did not believe in infinite space. He argues that a space is a place that can be occupied by a body, regardless of whether there is a body occupying it at the present. An infinite space would be susceptible to being occupied by an infinite body, which is impossible.

This schema made it necessary to conceive of the movement of planets and stars as perfectly circular, since they had continuous movements. If they had moved in a straight line instead, they would have required an infinite space to continue. This configuration of the cosmos was inherited by Copernicus, Tycho Brahe and even Kepler, who shared these ideas on space and infinity.

For the Eleatic school, of which Parmenides (530–460 BC) and Zeno (490–430 BC) were members, reality, or the universe, could not have an origin and, as such, a start or an end. In this respect Parmenides stated that: "All is one, immovable and infinite, since the limit would border on the void." This leads us to a dead end, since it means leaving behind the horror of infinity to fall into the horror of the void.

We have a series of concepts that we cannot comprehend but which are nonetheless there. There is not much difference between the horror of nothing and fear of the infinite. In fact one complements the other, although infinity often ends up winning the struggle, since, in a certain sense, it is closer to us. We cannot conceive of the space in which we live as being finite. When someone tries to imagine it in this way, the first question that springs to mind is: "What lies beyond?" The answer cannot be "nothing." There must at least be space, even though it is empty. The matter is simple. We do not know nothingness whereas we do have the constant presence of



There was no room for infinite space in Aristotelian thought. According to the Greek philosopher, an infinite space could only be occupied by an infinite body, which did not exist. This marble bust of Aristotle is a Roman copy of the original Greek statue in bronze by Lysippos dating from 330 BC.

things that make up infinity, even though it may be an imaginary infinity. Infinity is not just an idea or concept. Its presence, in any culture, together with the questions it raises, is an unmistakable sign, whether we like it or not, that it belongs to us, just like life, death, space and time.

Potential infinity and actual infinity

Imagine that we have drawn a line on the ground in chalk. If we were to take a step forward we would be on the other side of the line. This is an act that we can 'potentially' carry out. When we have done so and are on the other side of the line, we have 'actualised' this potential, converting it into a real act. There is a clear difference between what is potentially possible and an act that has been carried out. It may be the case, for example, that when we are on the verge of performing the action, we are struck by a sudden force that prevents us from crossing the line.

We say that the series of natural numbers 1, 2, 3, 4,... is infinite. In principle, nobody would doubt this, since given any number n we can always create another number that follows it, $n + 1$, regardless of how large the number n is. However, the possibility of doing this is one thing; actually doing it is another. The difference is subtle. The possibility of doing so defines potential infinity; having done so defines actual infinity. The choice of the words in English for designating these two classes of infinity is somewhat unfortunate, or at least not particularly intuitive. Perhaps the terms 'theoretical infinity' for potential infinity, and 'real infinity' for actual infinity would be better.

We know that nobody can construct the complete series of numbers. It is also the case that no one has ever seen two infinite parallel lines, since lines are infinite and at best we can see segments of parallel lines. Is this to say that parallel lines do not exist? They exist insofar as lines exist, but does an infinite line really exist? Euclid himself, in his famous *Elements of Geometry*, was extremely careful when discussing the matter, since when he refers to lines, he talks about "segments that can be made as long as we wish", in a clear reference to potential, but not actual, infinity.

The acceptance of temporal infinity, or actual infinity, is not just a problem of choice, but a philosophical position that is by no means trivial. It should be kept in mind that potential infinity was the only type of infinity permitted in mathematics (and science in general) until the end of the 19th century. Aristotle states the tacit prohibition of what would be later regarded as actual infinity in his school of philosophy: "The infinite cannot be an actual thing and a substance and principle," he wrote, adding, "It is clear that absolute negation of the infinite is a hypothesis that leads to impossible consequences," such that the infinite "exists potentially [...] that is to say, by addition or division."

Hence, the Aristotelian regulation of the infinite does not make it possible to consider a segment as a collection of aligned points, but it does allow this segment to be divided in half as many times as we wish.

INFINITY AND CHURCH PRIESTS

During the Middle Ages, the debate regarding actual infinity was not yet able to take on a mathematical nature, since it was an exclusive property of the divinity, and as such, could only be debated in theological forums. As St Augustine stated, "Only God and his thoughts are infinite." However, it is surprising that church priests deny God the possibility of believing in actual infinity. In his *Summa Theologiae* St Thomas Aquinas proves that although God is omnipotent, unlimited, he cannot create things that are absolutely unlimited. This conclusion, drawn by Aquinas in the context of religion, was justified through the admission that actual infinity is identical to pure evil.

Let's imagine the following questions on the idea of infinity were put to someone who lacked a philosophical or mathematical background. Their replies were quick, without thinking too deeply, answering spontaneously and guided by the 'common sense' of our cultural surroundings.

Q: What is infinity?

A: *Something that never ends.*

Q: What does that mean?

A: *You could be counting and never reach the end.*

Q: Why would you never reach the end?

A: *Because there is no final number.*

Q: How do you know this?

A: *I have not tried to verify it, I believe it.*

Q: Then we are talking about a belief.

A: *Not exactly, I know that no matter how large a number may be, a number can always be added to it.*

Q: I disagree. Even if you dedicate your life solely to this task, your life is limited and you cannot go on adding numbers forever.

A: *That doesn't matter, generations and generations of beings could dedicate themselves to the task.*

Q: Life on Earth is also limited. In fact, the Solar System has an expiry date.

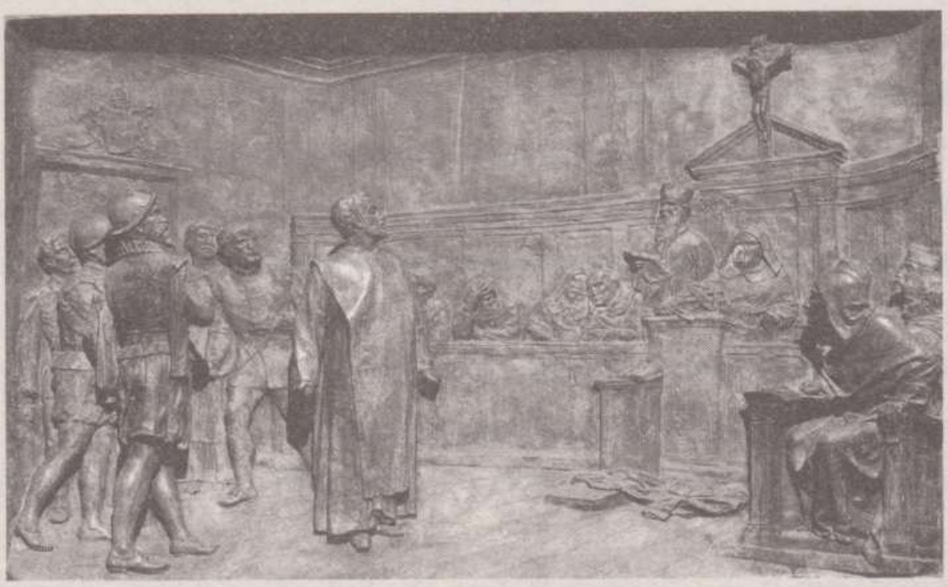
A: *That doesn't matter. It is not necessary for somebody to do it, it's enough to know that it can always be done. Even if there were only dolphins on Earth, it could be done. The fact that nobody can do something, does not mean to say that it cannot exist.*

Q: This is accepting that infinity is something that exists independently of us.
A: *Of course.*

This conversation gets to the crux of the debate on actual and potential infinity. The person asking the questions has clearly come down in favour of Aristotelian thinking.

THE FLAMES OF INFINITY

In 1600, Giordano Bruno (1548–1600) committed a 'thought crime' by imagining that we lived immersed in an infinite space populated by an infinite number of worlds. He then committed the error of expressing this thought publicly, which led to him being burned at the stake. Prior to this, he had been held in prison for seven years, subjected to all sorts of humiliation and torture. This treatment revealed two things: firstly, Giordano's firm conviction in his idea of infinity and the freedom of thought, and secondly, the danger of opposing the cultural milieu at certain points in history. In hindsight the Giordano's fate leaves a bitter taste: the modern scientific community has currently reached a consensus that the Universe could be infinite. The conclusion we can draw is that an idea is nothing more than an idea, for which we may sacrifice our prestige but not our life. It is not worth it.



Bas-relief in bronze by the Italian sculptor Ettore Ferrari (1848–1929) depicting the trial of Giordano Bruno by the Roman Inquisition. Campo de' Fiori, Roma.

Infinity in teaching

Potential infinity becomes part of our mental structures from the first years of schooling. It is closely related to the idea of counting objects and, as such, the series of natural numbers or cyclical processes related to the passing of time: day is followed by night, night by day, and so on. It is often a mental representation that does not change and which, if it does conflict with our intuition, does so without causing alarm. In fact, it remains more or less stable in the representations of our mind, because we do not make great requirements of it.

On the other hand, with actual infinity, things work differently: it always appears in a mathematical context, and does so without warning and without sufficient preparation. It is guaranteed to cause problems; in many cases it becomes a conflict that is very difficult to overcome. This conflict appears in its fullest form when we begin to study calculus. Research has been carried out (and continues to be carried out) to identify and evaluate these critical elements in teaching mathematics and, more specifically, calculus.

For the layperson, it should be pointed out that when we talk of calculus in this context, we are referring to the so-called infinitesimal calculus, a branch of mathematics that students often first encounter at around 16 years old in secondary school, and which continues throughout the majority of technical or scientific studies for two or three more years.

ACCEPTING ACTUAL INFINITY

The majority of surveys and studies carried out among the general public have shown that half of those interviewed do not accept the existence of actual infinity. What is interesting, however, is that it is not an issue of maturity, since analysis shows that the believers are not especially old or young. It is even the case that teachers, when forced with having to explain definitions or theorems in which actual infinity plays a determining role, 'play the game', while internally maintaining a fixed position with respect to the subject; quite simply, they believe actual infinity does not exist.

The inclusion of set theory in secondary teaching, which was accompanied by the false epithet of 'modern mathematics,' has been regarded by many in the teaching community as a complete disaster. Perhaps this is because the theory is of mathematical interest as the theoretical foundations of the discipline, but is of little practical value. As a result, in the majority of schools, teachers limited themselves to teaching very basic concepts, such as the membership of a set and inclusion among sets, concepts that are highly intuitive and do not require a mathematical language beyond their own symbols. On the other hand, an interesting area is avoided that makes reference, as we will see in the final chapters, to the concept of cardinality (the number of elements contained in a set), especially when the concept is applied to infinite sets.

In this context, we always speak of actual infinity, and it is here where students come into conflict with what they regard as 'common sense'. They will be required to accept sets of elements in which certain parts are equal to the whole. This is something that Euclid himself set out to clarify in his *Elements* by specifically stating that "the whole is greater than the part," as it appears should be the case if we follow correct logic. Similarly we find it hard to accept that a bounded set may still be infinite, since, according to our understanding, the infinite does not have limits.

As we shall see throughout the course of this book, elementary logic, which we refer to as 'intuition', can create misconceptions when dealing with actual infinity. Indeed, when dealing with certain concepts we often confuse understanding with believing. The bafflement that actual infinity provokes in maths students is similar to that caused by quantum mechanics in physics students. A typical example of the latter is as follows. Assume we have a box with a marble and two holes. If we move the box randomly, we can expect the marble to fall through one of the two holes. For certain movements, we can even accept the possibility of calculating the probability with which it will fall through one of the holes. However, it is more difficult to accept the marble falling through both holes at the same time. This possibility exists in quantum physics and clashes head on with our intuition. It is not a matter of understanding the phenomenon; everybody knows what the phrase "fall through two holes at the same time" means. Therefore in this situation it is more accurate to say "I don't believe it" instead of "I don't understand".

Something similar occurs with actual infinity. When we state that a small line segment contains an infinite number of points, we are aware of what we are saying, however whether we believe it or not is another matter.

ARCHIMEDES' SAND RECKONER

Our word for *millions*, which allows us to enumerate large quantities, was first introduced by the French mathematician Nicholas Chuquet (c. 1445–1488) in 1484 with the suffix *-illion.* $M = 10^6$ (according to this naming convention, M1 is a m-illion, M2 = bi-llion, M3 = tri-llion, etc.). Ancient numbering systems were often caught short when it came to handling extremely large numbers. In ancient Greece, numbering systems did not go beyond 100 million. Archimedes wrote a famous work on arithmetic, known as the *Arenarius* (meaning the Sand Reckoner), in which, amongst other things, he deals (theoretically) with the problem of counting the number of grains of sand on Earth. In fact, his aim was to prove it would be possible to establish a numbering system that was suitable for counting finite collections of objects that nonetheless appeared infinite. For doing so, Archimedes developed a system consisting of three periods based on the successive powers of myriads (Ω), a measure equivalent to 10,000 units. Using this system, the highest number he achieved was $10^{8 \cdot 10^{16}}$, a highly respectable number. What nobody has explained is why he stopped at that number, when there was nothing to prevent him from continuing.

Chapter 2

Discrete and Continuous

The opposing nature between the 'discrete' and the 'continuous', which has been a fascination of many thinkers, dates back to the philosophy of ancient Greece. It is now used in sciences as diverse as physics, mathematics, psychology and linguistics.

Density

The great cultures of antiquity, especially Greece, attached a metaphysical significance to numbers. Their world view depended on a numbering system. Generally speaking, when we talk of numbers, we are referring to the series of natural numbers, $1, 2, 3\ldots$, since the Greeks regarded fractions as proportions between quantities or ratios of similarity between given geometric figures rather than as numbers in their own right.

Before we go on, it is necessary to clarify an aspect of numerals directly related to the concept of infinity. If everything in existence needs to be explained using numbers, we would need enough numbers to designate each and every thing that is known and, furthermore, that may be known. As such, the series of natural numbers does not present problems, as it goes on forever. However, the series of fractions has a property that integers lack, and which Greek mathematicians observed with a certain degree of suspicion. This property is named 'density'.

There can be no integer between any two successive integers. For example, there is no natural number that fits between six and seven. Obviously, we would need a number that is greater than six and less than seven; impossible when confined to the natural numbers. However, it becomes entirely possible when we supplement the set of natural numbers with fractions. Returning to our example, this time the number

$$\frac{6+7}{2} = \frac{13}{2}$$

does lie between six and seven.

Using this system, it is always possible to find a number that falls between another two. For any numbers A and B, it is guaranteed that

$$A < \frac{A+B}{2} < B.$$

However, for this condition to hold, our collection of numbers must include fractions, also known as rational numbers.

As this process can be repeated indefinitely, we know for certain that between any two given rational numbers, there is always an infinite number of rational numbers. This is referred to as density and puts an end to the idea of the 'next'. In the set of natural numbers, we can say unambiguously that the number 13 comes after 12. However in the set of rational numbers, it is meaningless to talk of the number that comes after N. If we claim this number is M, we are always mistaken because

$$\frac{N+M}{2}$$

comes before M.

Density presents us with a perception of infinity to which we are normally unaccustomed. To take an example from geometry, when we imagine a straight line, we do so imagining that it extends indefinitely in both directions. This is our perception of the infinitely large. However, we can carry out the same procedure that we have carried out for fractions with the points of a line. It is always possible to find another point between two points, such that the number of points contained in a segment is also infinite.

Discrete and continuous

In the dictionary the meaning of 'discrete' is 'individually separate, distinct,' which refers us to the mathematical meaning of a 'discrete quantity': 'That which consists of units or parts separated from each other, such as the trees on a hill, soldiers in an army, grains of corn, etc.' As we shall see later, the reference to 'separate parts' defines the discrete by means of a highly advanced mathematical concept, since it should be noted that in mathematics the term 'separate' is not as obvious as it may seem.

To correctly understand the ins and outs of infinity, both in terms of the infinitely large and the infinitely small, we need a clear idea of the meaning of the concepts 'continuous' and 'discrete'. Let us consider a simple example that illustrates the

difference between the two. Imagine we have two equal receptacles, one of which contains water, and the other marbles. Pouring the first into a vessel, as we watch the liquid flow, we can see the level of the receptacle rising as it fills with water. If we add the marbles instead, the situation is completely different, especially in terms of our perception: the marbles drop one by one into the receptacle. These are two completely different experiences, not just visually, but also audibly: in the first example, we hear a continuous noise, whereas in the second, we can distinguish between the sounds made as each marble hits the receptacle. As we will have now already deduced, in case of the former, we are dealing with a continuous process, whereas the latter is discrete.

Let us consider another example: time flows continuously from nine o'clock in the morning until nine o'clock at night. However, if we consult a train timetable during this same interval, we find a set of discrete values. If one train leaves at 10 o'clock in the morning and the next at 11 o'clock, there is no value between 10 and 11, because these are discrete values. In contrast, the measure of time between 10 and 11 is a continuum that contains infinite values of which 10 hours 25 minutes and 0.3462876172004124474 seconds is just one example.

The way we have been discussing the concept, we might be forgiven for thinking it is intuitive and apparently simple. However, it has been the cause of impassioned conflicts throughout history, partly due to the fact that the matter is far from simple, and partly because, as we shall see later, intuition is not always good counsel, since the same object may appear both continuous and discrete, depending on the scale used to observe it.

The polemic surrounding the continuous–discrete antithesis is closely related to the concept of infinity, and there is nothing surprising about the fact that it is a discussion largely for the practitioners of philosophy – as it has been since the conflict in ancient Greece between the Pythagorean and Eleatic schools. Zeno's paradoxes serve as good examples and are coming up later.

The key question of whether the world in which we live is discrete or continuous comes down to scale and is therefore tangled up with our ability to perceive the very large and very small. Far removed from philosophical navel gazing and psychometric investigations, the hard sciences of physics and mathematics were thrust into the world of the discrete at the start of the 20th century. The former arrived there thanks to the appearance of quantum mechanics, and the latter as a result of the birth of so-called discrete mathematics.

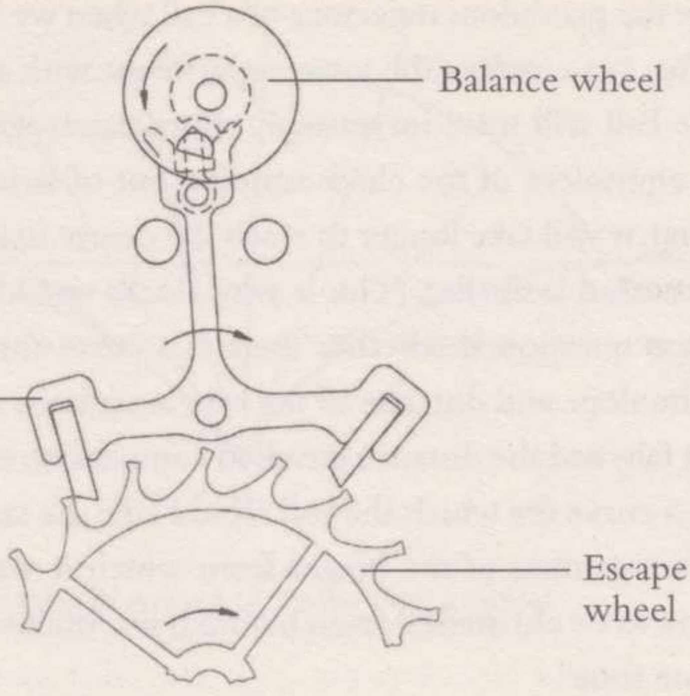
Trapping time

It is said that the most significant difference between science and technology is that the former changes how we see the world, whereas the latter alters how we live in it. In this respect, we can observe that the mechanical clock constitutes one of humankind's most revolutionary inventions, and one which, furthermore, has exerted the greatest influence on our everyday lives. Moreover, thanks to a device in which mathematics plays a decisive role, time was no longer observed as a continuous process and became conceived of as a discrete series of intervals.

The first mechanical clocks appeared in Europe in the 14th century (in China they date back to the 10th). They had mechanisms that would now be regarded as rudimentary. Their movement was based on a weight that descended under the force of gravity and hung from a chord wrapped around a cylinder. As the weight descended, the cylinder turned and moved the workings of the clock. As clock faces and hands had yet to be invented, the hours were marked by bells – one large clock would serve a whole community. In many languages, the word 'clock' refers to the idea of a bell (i.e. the French word *cloche*). The bells were rung manually by someone who watched the actions of the clock.

It goes without saying that the precision of these clocks left much to be desired, not only due to the imperfections of their workings, but also to a matter of elementary physics. The weight that powered the mechanisms did not descend at a uniform speed, since it was subject to the effects of gravity, meaning the speed of the weight would have increased as it descended.

An ingenious mechanical invention, known as an escapement, was largely responsible for solving the problem. The basic clock mechanism is made up of a toothed gear, an escapement and a pendulum. The escapement locks the escape wheel at one of its ends. When it is balanced, the device releases its lock only to apply it again immediately. The action of escapement – powered by the swing of a pendulum – produces the sound we now know as 'tick-tock,' which from that point onwards, has trapped, or at least regulated, most of our everyday lives.



This mechanical device, known as an escapement, improved the accuracy of timekeeping. The balance wheel makes a rhythmic oscillation that moves the pallet from side to side. With each oscillation, the pallet moves the escape wheel by a single tooth, regulating the movement of the clock mechanism.

However, this still left a serious problem to be solved: how could clocks mark a rhythm of time that (forgive the contradiction) would not depend on time? The problem was related to the fact that the first hours were longer than the last ones, or rather the clocks ran fast as the chord ran out due to the fact that the pendulums followed a circular path. There is an easy way of understanding this phenomenon if

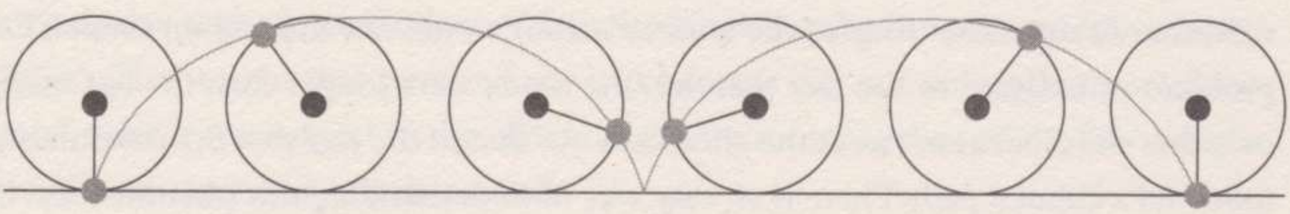
A CURIOUS TOY

Imagine we invert a cycloid (see page 30), place it on a table and rotate it. This would mark out a concave surface akin to asking a potter to make us a basin with a curve defined by a cycloid. An object with this shape was actually manufactured (in plastic) and sold in the 1960s in some curiosity shops in the United States. What is so special about this basin? If we roll a marble around inside, it always takes the same time to reach the bottom, regardless of the height from which it is dropped. It is fascinating to see how two marbles, one placed on the upper edge of the basin and the other halfway up the opposite side, both reach the lowest part at the same time.

we observe the pendulous trajectory of a ball when we let it go inside a semicircular surface. (You can conduct the same experiment with a marble and a suitable large bowl.) The ball will trace increasingly short trajectories until finally coming to a rest – the equivalent of the clock running out of string. It is clear that when the ball is higher, it will take longer to reach the centre than when it is lower, since the distance travelled is smaller. (This is why clocks would run faster, later.)

The next question is whether there is a curve for which this is not the case, whereby the slope and distance to the base are related in such a way that the speed at which it falls and the distance travelled compensate each other? The result would have to be a curve for which the ball would take the same time to reach the vertex of the base regardless of the height from which it was dropped. This feature led such a curve to be christened (even before it was discovered) a 'tautochrone', which means 'same time'.

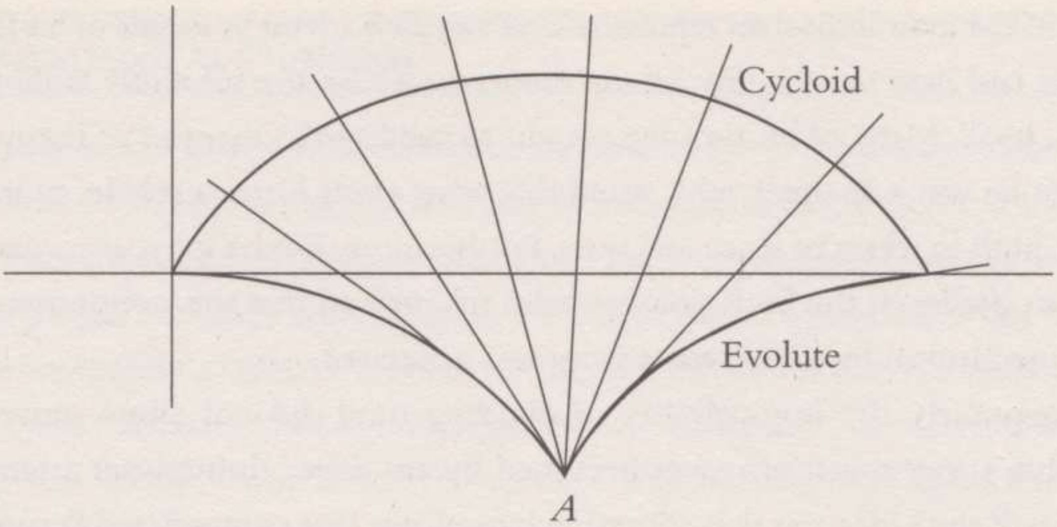
In 1673 Christiaan Huygens showed the 'cycloid' satisfied the properties of a tautochrone. The cycloid is defined by the path described by a point on a circle that was rolling along a straight line.



The drawing shows how a turning circle forms a curve called a cycloid.

Huygens had the idea that if a pendulum travelled in cycloidal motion, the height from which each oscillation began would not matter, since the situation would be the same as for the marble in the bowl, for which the time taken to reach the lowest part was always the same.

But how would it be possible for the trajectory followed by the pendulum to be cycloidal? The solution to this problem lies in one of the cycloid's most fascinating properties: 'the evolute of a cycloid is also a cycloid'. The idea of an 'evolute' is too complicated to be explained here, but we can see the geometric translation of this result. Imagine that we begin by splitting a cycloid in half and joining both halves at a vertex A, as in the following drawing:



Construction of an evolute derived from a cycloid.

If we take a fixed length of thread fastened to the vertex A and extend it such that it is always supported on one of the branches of the cycloid, the end of the thread traces a curve that is also a cycloid. Huygens had discovered a way to build a self-adjusting pendulum by turning the above drawing around, so that the movement of the pendulum would be constrained by the two branches of the cycloid.

Although time cannot be conceived of as a physical quantity in the way that we can measure mass or temperature, based on Huygens' invention we are able to state that it can be conveniently handled as a discrete quantity.

Our everyday life continues to fit the beat of the 'tick-tock,' a discrete measure of time. However, in the realm of science, the interval between the tick and the tock has been shrinking in an astonishing manner. It is infinitely smaller, colloquially speaking, than a second. Modern atomic clocks mark cycles in which a second is divided into 9,192,631,770 parts, making the clocks incredibly discrete!

Zeno's paradoxes

The discrete is made up of elements, of units, but what about the continuous? It seems logical to think that the continuous may not have this structure, since units are separable and there can be nothing between two units that are in contact, since if there was something, it could also be divided into units. As soon as we begin to reflect on the matter, we begin to see that the concept of the infinitely small leads us directly to continuity. From the outset, reflections on the nature of the continuous played a significant role in Greek thought.

One of the most important representatives was Zeno, who by means of his famous paradoxes laid bare the fragility of any theory based on the infinitely large or the infinitely small. Many of his writings sought to validate the theories of Parmenides (of whom he was a disciple), who stated that what exists forms a whole, an indivisible unit, both in terms of space and time. Furthermore, Zeno's spirit also contained a desire to challenge the Pythagoreans, who maintained that the 'continuous flow' was the mechanism by which everything was generated.

Consequently, the impossibility of dividing time did not allow movement, regarded as a succession of spaces occupied by an object throughout a temporal succession. Zeno's idea was that admitting hypotheses that contradicted Parmenides would cause contradictions to occur, absurdities that could not be tolerated by reason. To do so, he made use of a logical method of which he can be regarded as, if not the creator, at least an early champion: the so-called method of *reductio ad absurdum*.

Essentially, the method works as follows: We assume a given hypothesis is true and use this as the starting point for a series of logical deductions, which lead to a result that is clearly false. We can conclude that the hypothesis was also false. In terms of logical statements, the schema is based on

$$\mathrm{T} \Rightarrow \mathrm{T}$$
$$\mathrm{F} \Rightarrow \mathrm{F}$$
$$\mathrm{F} \Rightarrow \mathrm{T}$$

where T = TRUE, F = FALSE and $\Rightarrow$ is an operator that means 'implies'. Hence, for example, $\mathrm{V} \Rightarrow \mathrm{V}$ means that one truth leads to another truth. This is to say that something that is true can never imply something that is false, or put another way: something that is false can never be deduced from something that is true, and if this is the case, it is because the starting point is not true. The *reductio ad absurdum* method is based on this law and as such is a form of reasoning whose aim is to demonstrate that a statement is false. This is the logical outline on which the paradoxes are based.

The Pythagoreans had a mathematical and geometric outlook that conceived of reality as being made up of points. The points created lines, lines created surfaces, and surfaces created three-dimensional bodies. Zeno argued against this, deducing that if points lacked dimensions, their magnitude could not be measured and, as such, nothing built using points could have magnitude either, and hence they had no range of existence. Furthermore, everything made up of points was susceptible to being divided an infinite number of times, leading to various absurd situations.

PARADOXICAL THINKING

A paradox is an argument that conflicts with common sense. Its aim is to cast doubt on an established principle by arguing from a position that produces contradictory conclusions to it. More precisely, logical paradoxes (of which the Eleatic school was a pioneer) are based on using logical statements to deduce others that may be indiscriminately true or false. One of the most popular paradoxes is known as the 'Liar's Paradox' and was proposed by Epimenides of Crete. It states the following: "All inhabitants of Crete are liars." Epimenides cannot tell the truth, since he is from Crete, but nor can he lie, since this would be to confirm something that is true, and therein lies the contradiction.

The paradoxes have impeccable logical architecture. They have been ruminated upon all the way down to our times and have many possible interpretations. They are a fundamental aspect of understanding the problem of infinity in all its dimensions. It is believed Zeno developed more than 40 paradoxes based around this theme, although of the ones that have been handed down to our times, the four best known are the dichotomy, Achilles and the tortoise, the flying arrow, and the stadium.

The dichotomy

This paradox directly attacks the notion of movement, showing it is impossible based on the fact that a body that moves between two points *A* and *B* will first have to cover the first half of the distance separating the points *A* and *B*. Once it has covered this distance, it will then need to cover the remaining half of the distance, and so on. This infinity of distances the object needs to cover cannot be covered in a finite interval of time. Hence, motion is impossible.

Achilles and the tortoise

Achilles (he of the heel) believed that man was faster (in terms of speed) than a tortoise. This paradox involves a race between the two. If the two set off at the same time, Achilles will obviously win. The trick consists in giving the tortoise a head start, however small as this might be. Under these conditions, Achilles must first reach the starting point of the tortoise. At that point, when Achilles arrives, the

tortoise, which has not stopped at any point, will no longer be there, but will be farther forward, regardless of how small the distance. This means that Achilles will be required to cover another distance that separates him from this point. However, when he arrives, the tortoise will have moved on again, and Achilles will not have reached the point. As this process can be repeated *ad infinitum*, Achilles will never manage to catch up with the tortoise.

Although it may seem that both these paradoxes are, if not identical, extremely similar, there is a subtle difference between them. In the case of the former, space is always divided into equal segments. However in the paradox of Achilles and the tortoise, the space is divided in decreasing proportions proportional to the speeds of the runners.

The flying arrow

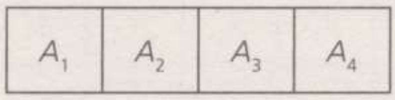
This is the most confusing paradox of the four. Historians have deduced that the original text was corrupted in some way and have had to recompose it. It states that when we launch an arrow into space, we feel like it is moving away from us, whereas in reality, it is not moving at all, since the arrow, like any other object, must occupy a space that is equal to itself, and hence must be stationary. If time is made up of indivisible instants, the arrow cannot occupy two or more places at the same time.

Hence, just as the first two paradoxes refer to the impossibility of infinitely dividing space, this is based on the indivisibility of time, and, more specifically, the existence of what we call 'an instant', since if it is indivisible it lacks duration, and hence there is no motion. Thus conceived, the mental construct of an instant is the same as the point in geometry – and equally flawed according to Zeno.

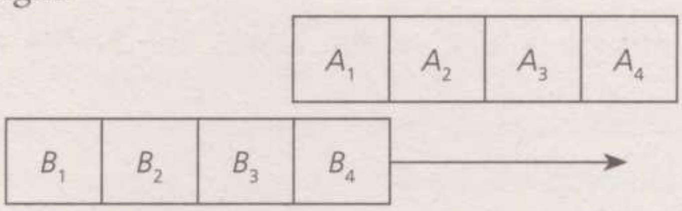
The stadium

Let us imagine that time is a discrete magnitude that we can imagine as small as we wish, but with a fundamental unit τ. Hence there is no unit of time smaller than τ, which is, as such, indivisible. We can imagine a clock in which each 'tick' or 'tock' corresponds to one of these indivisible units.

Let us now consider four equal bodies, A_1, A_2, A_3 and A_4, in a state of rest (originally the paradox uses a rank of four soldiers):

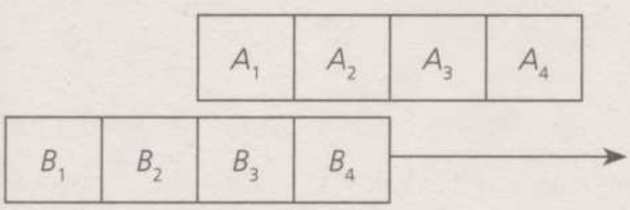


and four other bodies, B_1, B_2, B_3 and B_4, identical to the previous ones and which move towards the right:

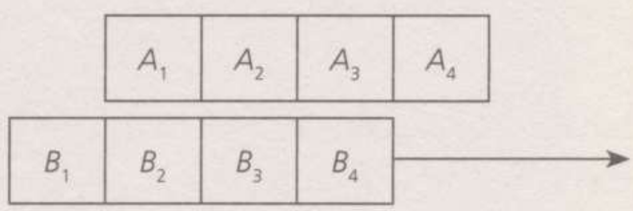


but do so in such a way that at any instant in time one of the bodies from B moves past another from A:

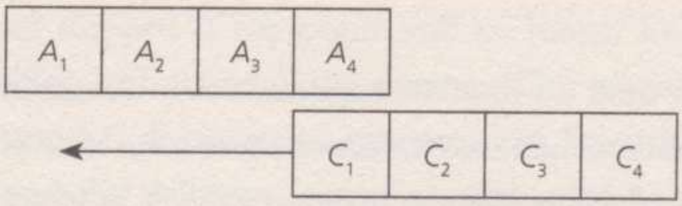
'tick'



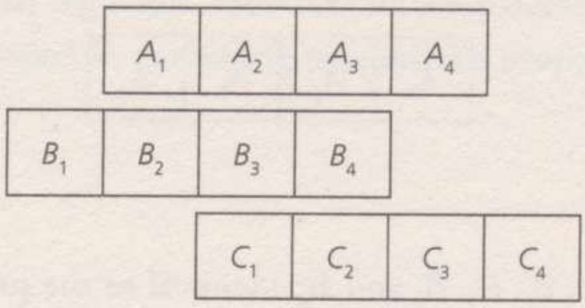
'tock'



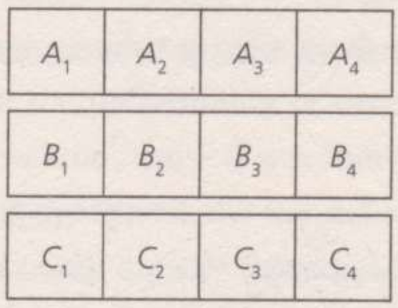
Now let us consider a third series of bodies C_1, C_2, C_3 and C_4, also identical to the previous ones, but moving towards the left in such a way that at any instant in time each of these moves past of one of the bodies A to the left:



The paradox arises when we consider the two movements of the bodies *B* and *C* at the same time. If we start from the following relative position:



in the next time interval (a 'tick' of the clock), the bodies are in the following position:



but this assumes that C_1 will have overtaken two of the bodies *B*; and hence we can divide the time into two which, of course, contradicts the hypothesis that it was an indivisible unit.

This is the most controversial of the four paradoxes presented here, partly because Aristotle demolished it by showing that Zeno referenced the bodies at rest and the bodies in motion in the same way. If the speed of the body in motion is uniform, the speed at which it passes another body in motion cannot be regarded as equal to the speed at which it passes another at rest. The paradox is regarded as controversial insofar as Aristotle's critique appears rather trivial but devastating and it is hard to believe that Zeno was not aware of it himself.

Other interpretations make amends by suggesting the possibility that the paradox would be more focused on the fact that, as in the previous examples, motion and time can be divided infinitely, and hence for one of the objects to be able to exceed another in motion, it would first have to pass the halfway mark of the one at rest, and so on. However, it seems quite plausible that Zeno, once again, wished to confront the Pythagorean school with a contradiction in terms of the indivisibility of geometric elements.

ZENO: A FORGOTTEN GENIUS

Zeno of Elea (c. 490–425 BC) was a Greek philosopher of the Eleatic school founded by Parmenides. Zeno's work has reached us through *Parmenides*, one of Plato's *Dialogues*. His philosophy can be summarised as monism, which, in a few words, states that all is one and change does not exist. In many respects, Zeno has not received the recognition that he deserves, according to some philosophers. Bertrand Russell partly corrected this omission by saying: "In this world, nothing is more capricious than posthumous fame. One of the most notable examples of posterity's lack of judgement is the Eleatic Zeno. This man, who may be regarded as the founder of the philosophy of infinity, appears in Plato's *Parmenides* in the privileged position of instructor to Socrates. He invented four arguments, all immeasurably subtle and profound, to prove that motion is impossible; Achilles can never overtake a tortoise, and that an arrow in flight is really at rest. After being refuted by Aristotle, and by every subsequent philosopher from that day to our own, these arguments were then reinstated, and made the basis of a mathematical renaissance..." (*The Principles of Mathematics I,* 1903).

A fresco from the library of the El Escorial monastery in Madrid depicts Zeno of Elea showing his disciples to the doors of truth (veritas) *and falsehood* (falsitas).

On the other hand, Aristotle's critique of the first paradox left us the foundations of a concept that would become essential in the history of infinity; something also regarded by many writers to be one of the most important contributions ever made on the concept of infinity.

Firstly, it draws our attention to the fact that the word 'infinite' allows two interpretations: infinite in terms of extension and infinite in terms of divisibility, and that both concepts are confused in this paradox, since upon being applied to time and space, they ensure a limited space can be covered in a finite period of time, regardless of whether it is infinitely divisible. It then makes the following distinction: the continuous space in which the body moves has an infinite number of halves, but only potentially, not actually. Herein lies the significance of Aristotle's contribution, since from this point onwards, when the term 'infinity' was used, it was done so in two extremely different, and to a certain extent irreconcilable, ways: the so-called potential infinity and actual infinity that we saw in the previous chapter.

We often establish what is true or certain by means of common sense, which in the end, common or uncommon, is based on what we call the senses, which can be defined in terms of modern technology as the devices through which we perceive and process the reality that surrounds us. Something is reasonable insofar as our perception shows us how it is manifested. As paradoxical as the arrow's flight may seem, the perception of our senses permits no other interpretation – the arrow moves away from us.

Zeno was doubtless acutely aware of this fact, although he also knew that on certain occasions our senses may not be a reliable basis for reason, leading him to formulate the following argument. In the same way that something does or does not have magnitude, an object does or does not produce sound. A basket filled with grains of millet makes a certain sound when we pour it out on a surface. This led Zeno to ask if a single grain would make a sound. If this is the case, the next question is whether half a grain of millet would produce a sound. As we can imagine, the process is one whereby the successive division of the grains will lead us to a point at which the sound will no longer be perceptible. Starting from this point, it is possible to state that the sum of elements whose value is zero will always be zero. If we combine many things that do not produce sound, neither will the end result produce a sound. Zeno's aim is to show that we cannot trust our senses when we enter into a certain realm of reasoning. This is a school of thought in which the senses give rise to intuition, especially in mathematical reasoning, but, as we shall

see later when we come to Cantor's theories, intuition can also deceive us, and we cannot use it as the basis for our reasoning if we wish to consider areas of thought in which infinity is real and an object that we can handle with the same ease as if it were any natural number.

Zeno argued that if the unit is made up of infinite parts, this is only possible if each does not have a magnitude, since if this were the case it could be divided and hence would not constitute a unit. However, this would mean the object constituted by the units would also lack magnitude, since the sum of elements that do not have magnitude cannot have a magnitude either. This was how the Greeks defined the word *apeiron* in place of 'infinity'. In terms of magnitude, to be *apeira* means to lack a defined limit, a concept more consistent with the idea that an object is infinite because it can be made as large as we wish. The idea is not so much to return to the example of numerical series, that the numbers never reach an end, but that given any number, it is always possible to add another. The case is similar for the infinitely small, in the sense that it is possible to make things as small as we wish, a concept that in the analysis developed in the 19th century would be formalised by a rigorously established mathematical definition.

Squaring the circle

Diagrams constructed using a ruler and compass were part of the geometric problems of ancient time, and most thoroughly espoused by the Greeks. The variety of problems that can be set out ranges from extremely simple cases, through to more difficult ones, all the way to problems that are extremely difficult or even impossible. The most famous examples are squaring the circle, trisecting an angle and doubling a cube. Of all these, the first remains the most familiar, as reflected in the expression, when something present great difficulty, that: "This is harder than squaring the circle."

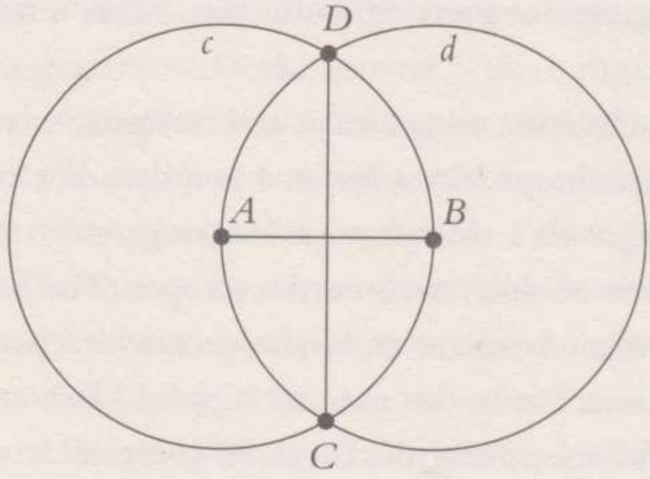
When we talk of constructs using a ruler and compass, we must adhere to very specific rules, otherwise the problems become pointless. For example, to find the midpoint of a segment with a ruler that has markings every millimetre is, to put it simply, cheating, since we don't even need a compass. The first step is to specify what we mean by a 'ruler'. A ruler is an ideal object, with a perfectly straight edge for drawing straight lines, but in this exercise it should lack any type of marking that could be used for measuring distances. A 'compass' is a normal, standard, movable compass, in the sense that it can be opened and closed at any angle. It

cannot be used to make marks that may later be used as distances, since this would be the same as having a ruler with distance markings.

MASCHERONI'S COMPASS

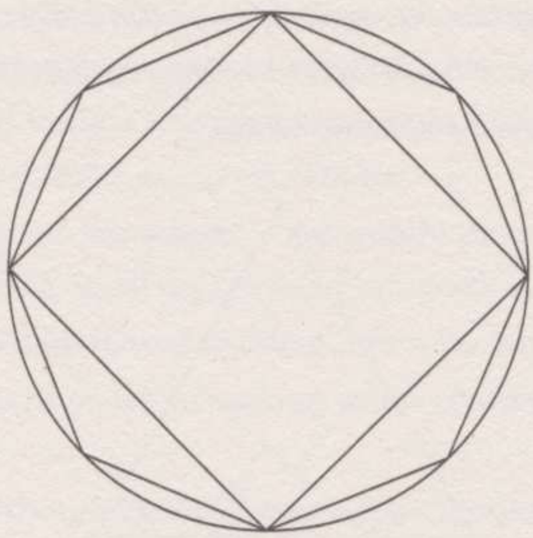
Historically, constructing geometric forms using only a ruler and compass have occupied a special place in recreational mathematics. One of the most curious publications in this field was by William Leybourn, a surveyor who in 1694 published a book entitled *Pleasure with Profit*. This tome described all sorts of mathematical games using a ruler and 'fork'. At the time, forks were made of parts that could also serve as a fixed compass. However, the great revolution in the history of constructions using a ruler and compass occurred in 1794, when the Italian geometer, Lorenzo Mascheroni, published his *Geometria del Compasso* with a proof that any construction made using a ruler and compass could also be made using only a compass (a movable compass, of course). In this respect, since it is impossible to draw straight lines using a compass, straight lines are defined by two points given by the intersection of arcs.

Having defined the rules of the game, we can now look at how to solve a problem. For example, let's see how to trace a perpendicular line through its midpoint. Imagine the segment has two ends, *A* and *B*. First of all, we must draw a circumference with centre *A* and radius *AB*. We then draw another circle with the same radius, only with centre *B*. The straight line joining the points where both circumferences intersect is a perpendicular line that passes through the midpoint of line AB.



Before going any further, we should note that trying to square the circle using a ruler and compass is futile, since the German mathematician Ferdinand von Lindemman (1852–1939) proved in 1882 that π was a transcendent number, and hence there is no solution to the problem of squaring the circle.

What has been proven, however, is that it is possible to construct any regular polygon whose area is equal to a given square. Although this has been theoretically proven, squaring a regular polygon is not always easy. Based on this result, Antiphon of Athens (c. 480–411 BC) designed a system for squaring the circle, the logic of which is hard to refute. Antiphon's idea is as follows: start from the fact that it is possible to construct a square whose area is equal to that of a series of regular polygons that we shall describe below. For a given circumference, we inscribe a hexagon:



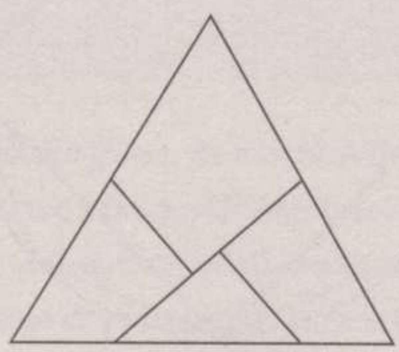
We already know it is possible to square the hexagon, that is to say, we can use a ruler and compass to construct a square that has the same area as the hexagon. The process now consists of increasing the number of sides of the polygon inscribed within the circumference, since we know we can square any such polygon. The difference between the area of any of these inscribed polygons and the area of the circle becomes ever smaller. In fact, we can make it as small as we like. Consider, for example a polygon with a quadrillion sides. Any of these sides will be extremely close to the curve of the circumference, to the extent that it may be impossible to distinguish between the straight line segment and curve. Antiphon believed it was possible to use this method to square the circle.

From a logical point of view, his argument is impeccable. However the only problem is that it entails a magic leap that takes place in the remote and inaccessible depths of the infinitely small, a place to which we do not have access.

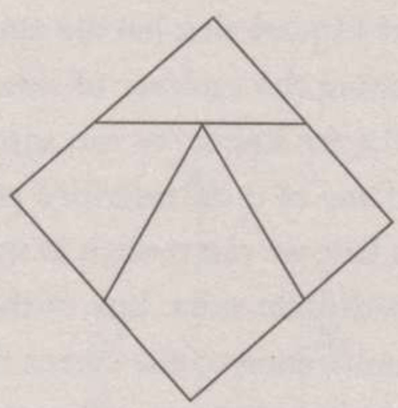
The circle is real, as are polygons with infinite numbers of sides. However when we come to look at the process that takes us from the infinite number of sides to a true circumference we enter the realm of actual infinity. Prior to that, we remain in the realm of potential infinity.

SQUARING A TABLE

Quadratures (squaring another geometric form) are often a difficult exercise, even for the basic shapes, such as the triangle, the pentagon and the hexagon. Some of the solutions with a compass and straight edge are even signed by their proud authors. For example, to square an equilateral triangle, we must follow a process whereby the triangle is decomposed (always using a ruler and compass) in the following manner.



These pieces can be used to construct a square that has the same area as the triangle:



Maty Grünberg took advantage of this geometric construction to design a table that can be used in the shape of a square or a triangle, depending on the situation.

Irrational numbers

The numbers 1, 2, 3,..., normally used for counting, are essential for measuring things. If we take a more or less straight piece of wood and mark a line for each number, such that all the numbers are the same distance apart, we can begin to measure, or establish lengths. The distance between two numbers is what we call a unit of measurement.

Let us assume that our unit is determined by a segment OA and that we wish to measure the length of a bar B. We place our unit on it and count how many times the segment is contained inside the bar. Let us assume the result is 5. In this case we say the bar measures 5 units. This time, we have been lucky, since the result is a whole number.

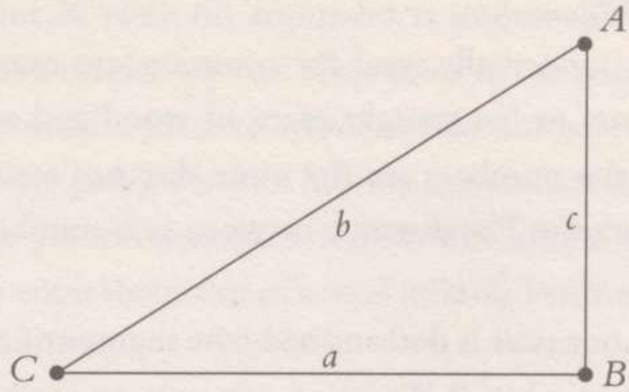
However, this is not always the case. Imagine, for example if the bar measured four and a half units. All we need to do though, is add another division to our units of measurement for representing a half. In symbols, this can be represented using a fraction. This is how we build a ruler for measuring, and the more divisions we make, the more precise our measurements will be.

It is clear that this process has a limit – the purely physical fact that relates the thickness of the marks to our visual capacity to distinguish between them. A normal ruler, such as the ones used by school children, often has millimetre divisions, such that the unit of measurement (if we are talking about centimetres) has been divided into ten.

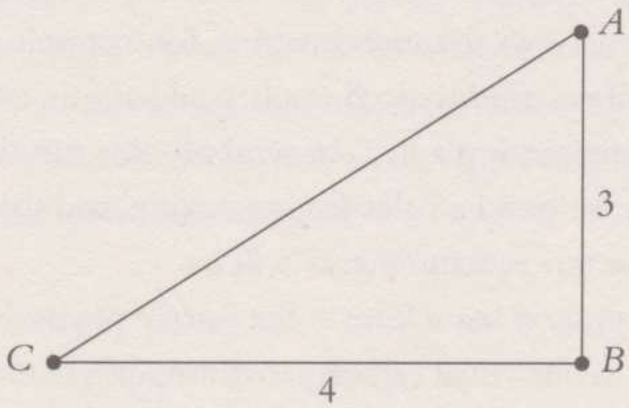
Before going any further, let's recall some basic concepts from geometry. A right-angled triangle is a triangle that has one of its angles equal to 90°. For example, the triangle ABC of the drawing on the following page is right angled, since the angle B measures 90°. The two sides that make up the rectangle are called the legs, and the third is the hypotenuse. Consequently, the hypotenuse is always the longest side and always faces the right angle.

The famous Pythagoras' theorem states that the sum of the squares of the legs is equal to the square of the hypotenuse. Hence, the following equation holds:

$$\overline{AB}^2 + \overline{BC}^2 = \overline{AC}^2.$$



This allows us to find the value of the hypotenuse, given the length of both legs. For example, in the following triangle

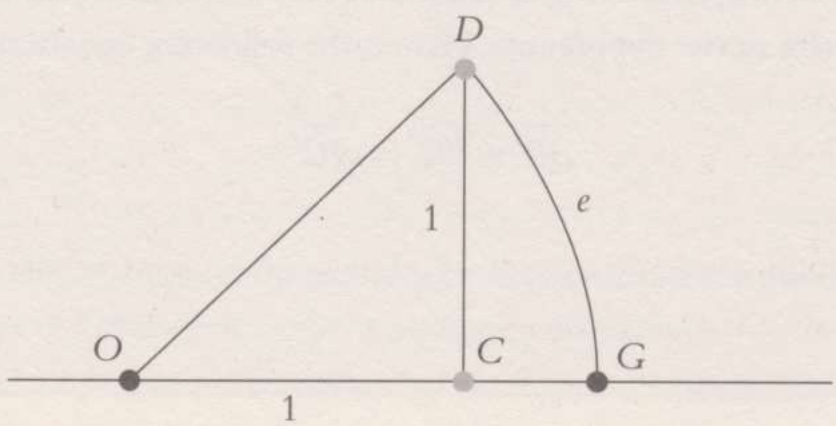


we know that

$$3^2 + 4^2 = \overline{AC}^2,$$

and hence the value of the hypotenuse will be 5.

Let us now assume we take a unit of measurement on a straight line, with origin O, such that $OC = 1$. Let us now draw a perpendicular segment through the point C, such that CD also has length 1. As can be seen in the following drawing, we have obtained a right-angled triangle OCD with hypotenuse OD.



Applying Pythagoras' theorem, we have

$$\overline{OC}^2 + \overline{CD}^2 = \overline{OD}^2.$$

Such that $\overline{OD}^2 = 1+1=2$, giving $OD = \sqrt{2}$.

If we use a compass to join the end of OD to the straight line, it is not possible to assign a known value to the segment OG. As such, we say that the length OG is 'incommensurable'.

Implicitly, what we are saying is that $\sqrt{2}$ cannot be written as a fraction, which leads us to a precise definition of a rational number. A number N is said to be rational only when it can be written as the quotient of two integers.

According to this definition, the following numbers are all rational: 2/3, 8/5, 2,773/12,452. Logically, integers are also rational, since any integer can be written as the quotient of another two (for example, 8 is the same as 16/2).

In certain apocryphal texts of Euclid's *Elements* we can find the proof that $\sqrt{2}$ is not rational. (A proof is given using modern language in the appendix.)

Numbers that are not rational are referred to as *irrational*, a description that aptly captures the nature of these numbers. However, the most serious aspect of this is that it is not only the diagonals of all squares that are irrational, but also between the height and the side of an equilateral triangle, and between the diagonal and the side of a regular pentagon. Hence, we have not only discovered the existence of one irrational number but something much more important: the existence of a set of irrational numbers. Integers cannot be used to precisely measure the most famous objects from Pythagorean thought. As a result, the discovery of irrational numbers created an unprecedented crisis in the history of Greek mathematics. Remember that Pythagorean sects were characterised by, among other things, a strong sense of secrecy preserved under oath. Perhaps the best-kept secret of all was the existence of irrational numbers, which contradicted much of the sect's philosophy. Legend has it that revealing their existence was punishable by death.

If we observe the numbers in decimal form, it becomes clear that there is a substantial difference between rational and irrational numbers. For example, the number ½ is expressed as 0.5 in decimal notation. In contrast, 1/3 = 0.333333333... has an infinite number of decimal places, which are nonetheless perfectly controlled, since they are always the digit 3.

A number such as

$$\frac{325}{100}=3.25$$

has only two decimal places. However,

$$\frac{95}{99}=0.4545...$$

has an infinite number of decimal places, although 45 is repeated infinitely. (We refer to the repeated number phrase as the *period*).

$$\frac{47113}{9000}=5.2347777...$$

is another type of decimal number in which the period appears after a non-periodic phrase.

The square root of two, on the other hand, is an infinite decimal expression in which the numbers appear without order or arrangement, randomly, as if decided by spinning a roulette wheel. Their formation cannot be explained by any law, or anything approaching it. Can we really say we know the value of $\sqrt{2}$? Regardless, if the answer is yes, we can only say for sure that we know approximate values, and that the approximation can be as precise as we wish, but nothing more… and nothing less, since when we say "as precise as we wish", we are implicitly saying that we have the power of exercising some form of control over the number's infinite series of digits.

The British mathematician Brook Taylor (1685–1731) established an approximation to $\sqrt{2}$ using the following series of numbers:

$$1,1+\frac{1}{2},1+\frac{1}{2+\frac{1}{2}},1+\frac{1}{2+\frac{1}{2+\frac{1}{2}}},1+\frac{1}{2+\frac{1}{2+\frac{1}{2+\frac{1}{2}}}},...$$

For each term in the series, we obtain a value that approximates to $\sqrt{2}$, with the peculiar feature that the numbers alternate between being to the left and right, as can be seen in the following table of values that gives the first nine terms.

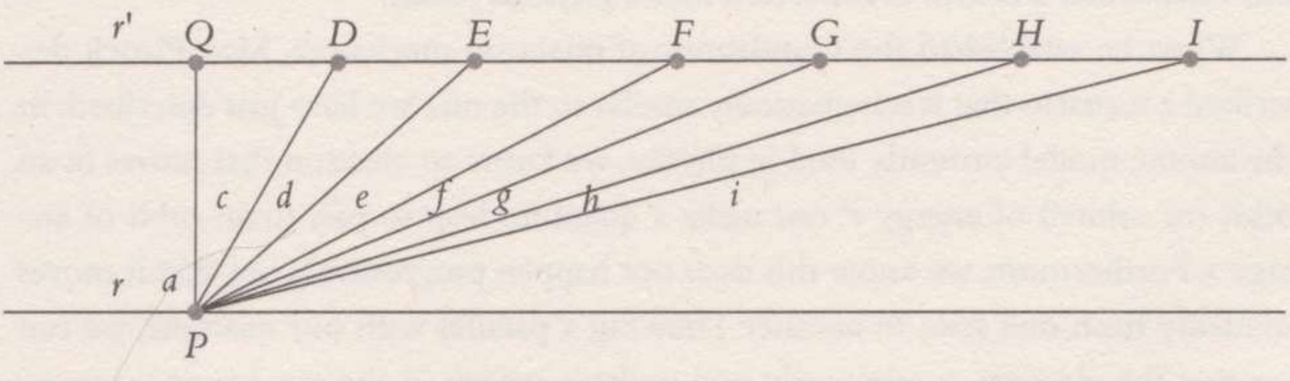
1	
	1.5
1.4	
	1.4166666
1.41379310	
	1.41422857
1.41420118	
	1.41421568
1.41421319	

Hence, starting from one, to the left of $\sqrt{2}$, and 1.5, to the right, we gradually get closer to the theoretical real value of the number. However, we are dealing with an infinite series that converges on the real value of $\sqrt{2}$. It is certain that this could well be the case, but affirming that $\sqrt{2}$ is a specific number is similar to accepting actual infinity.

If someone claims that irrational numbers do not exist, as the Greeks and many other mathematical figures maintained throughout the course of history, we can be sure that they are, albeit implicitly, denying the existence of actual infinity.

The quantum leap

Let us now look as a way of relating the infinitely large (infinite extension) with the infinitely small (the process of infinite division). Imagine we have two parallel lines, *r* and *r'*.



We fix a point *P* on the first, which we will use as a reference. Let's now take a second point Q, positioned, for example, perpendicular to *r* passing through the point *P*. The angle *a* that forms the segment *PQ* and *r'* is 90° (a right angle). We now move the point Q, located on the line *r'* to the right. Note that the angle *a*

changes, and as it does so it reduces the extent to which it is open as Q moves towards the right. It is clear that as Q moves farther away, the angle *a* will decrease. In this diagram, we have a clear link between the infinite extension caused by moving the point Q and the continuous reduction of the angle, which can be made as small as we wish. Expressed in a less technical way, we can say that as one element becomes infinitely large, the other becomes infinitely small. The important thing about this system is the following: point Q is moved towards the right of line *r'* in a continuous manner, just as the size of the angle also shrinks in a continuous manner.

We will now examine things from another point of view. The mechanism we use is to continuously reduce the angle and see how point Q moves away to infinity. The distance from point Q to the line *r* remains constant, and is the same as the distance between the two parallel lines. The key question now is what will happen when the angle formed by *PQ* and the straight line *r* is zero? The answer is that the point Q will have turned into a point at infinity, and not just at any point, but the point at which both lines converge. So far, so good, but the leap into the infinite has once again been counterintuitive. The potential infinity we have in mind becomes an actual infinity, with a surprising result: the distance from point Q to the line *r* has all of a sudden become zero.

Are we dealing with, as some of infinity's great thinkers have claimed, a process that can only be considered valid within a purely intellectual context? We shall never see the movement at which the point Q becomes part of the line *r*, and we are sure that when we 'actualise' this continuous movement towards infinity a new scenario will appear. In fact, modern physics provides us with an example in which this intellectual scenario is converted into a physical reality.

When he established the foundations of quantum mechanics, Max Planck described a scenario that is schematically similar to the one we have just described. In the atomic model currently used in physics, we know an electron that moves in an orbit (or orbital) of energy *r'* can make a quantum leap to pass to an orbit of energy *r*. Furthermore, we know this does not happen progressively but that it moves suddenly from one state to another. Drawing a parallel with our example, we can say that the electron continuously accumulates energy in the same way as we reduced the size of angle *a*. At a given moment, in the most accurate sense of this colloquial expression, the electron (our point Q) passes from one level of energy to another. As such, we can conclude that Zeno was correct. However this leads to a contradiction. There is no movement, according to our understanding of the word,

that takes the electron from one orbit to another, but there are simply two different physical states, something which, conceptually, occurs in that mysterious and fascinating realm in which potential and actual infinity live together in the absence of space and time.

Chapter 3

Encounters with the Infinite

The first people to allow us to 'see' infinity in spatial representations were not philosophers or geometers, but the artists of the Renaissance. Free from the stringent limitations imposed on their work by the Church, and thanks to the rediscovery of mathematical knowledge from the Greeks, they helped forge a new branch of mathematics, in which infinity was no longer stigmatised as a representation of evil.

Three-dimensional painting

When the word Renaissance is mentioned, numerous images of works of art come to mind, including sculptures, painting and architecture, in addition to certain technological advances. However, very little is said about mathematics. The period is known for the rediscovery of knowledge. The Dark Ages had consigned to oblivion (or the shelves of a few monastic libraries) the work of the Greeks and Arabs, which constituted the fundamental pillars of geometry and algebra. However, it was precisely in the field of geometry where the artists of the Renaissance carried out their most significant work, especially the painters. This is the métier where a geometric concept of infinity was developed.

In general, the artists of the Renaissance saw themselves as being required to develop different types of knowledge and skills, not just in art itself, but also in the sciences. It was often the case that their work was funded by wealthy patrons or princes who, in addition to paintings, would have many more demands, commissioning them to produce sculptures, works of music, buildings and fortifications to defend their properties – and even detailed mathematical analyses on the trajectories of projectiles that might threaten them.

At the beginning of the Renaissance, the artists had inherited a quintessentially religious conception of painting, characterised by a set of well-defined rules, relating to colour and shape. Figures of a holy nature, which were the majority, had to appear on golden backgrounds, symbolising that they inhabited the heavenly plane. Most colours were symbolic of a heavenly and earthly hierarchy, while position and size also worked to this effect. However, the most important thing was that all the

representations took place in an unmistakable two-dimensional space. Everything – and everyone – was shown flat, in a style barely developed since the days of ancient Egypt. This was not because medieval artists were just bad, but was rather an intentional subjugation to the symbolic order: sacred figures could not have a realistic representation in case that inferred certain earthly connotations.

THE SPIRIT OF THE RENAISSANCE

In his *Treatise on Painting*, Leonardo da Vinci (1452–1519), a typical example of a Renaissance genius (if there is such a thing), wrote a short passage in which he reflects on the concept of continuity. It is a highly philosophical text, not only on account of the idea it expresses, but also, in the true spirit of the Renaissance, on account of the number of disciplines it interconnects: "If you, a musician, tell me that only the sciences of the mind are not mechanical, I will reply that painting is of the mind, and that, like geometry and music, it deals with the proportions of continuous quantities and arithmetic, i.e. discontinuous ones, and that it deals with all continuous quantities and the qualities of proportions, shadows, light and distances, according to perspective."

Freed from the restrictions imposed by religious institutions, Renaissance artists were the first to represent reality as faithfully as possible. Put another way, reality took place on a three-dimensional stage and this meant the era saw the development of new drawing and painting techniques making it possible for the viewer to feel the sensation of spatial depth through the interplay of shadows and colours. Shadows, for example, indicated the position of objects, whereas the intensity of colours reduced as they moved further away from the foreground. These were all techniques that helped to recreate a sense of space. However, what was fundamental, essential even, was that the original drawing, the sketch, conformed to precise geometric rules. When this is taken into account, it is not so strange that the greatest mathematical advances of the time appeared through the art of painting.

More importantly still for our purposes, it was these artists who included infinity in their representations, turning what had until then been something that was merely a potential into a geometric reality. Remember that Aristotle said a line was only potentially infinite. Euclid improved this by defining a line as a segment that could be

extended as long as one wished in order to carry out any type of geometric construct or proof – and thus would it remain for all geometers until the 17th century.

However, during the 15th century, a position named the 'central vanishing point' appeared on painted canvases and architects' plans, giving rise to what would become known as 'central perspective'. This point, at which parallel lines meet, can be described as a point of actual infinity. Thanks to this type of perspective, artists such as Leon Battista Alberti (1404–1472), Filippo Brunelleschi (1377–1446) and Piero della Francesca (1416–1492), who reconstituted the theories of Greek geometry, managed to give viewers a clear sensation that a three-dimensional setting was being represented on a plane.

From perspective to projective

Has anyone ever seen two parallel straight lines? We can answer this question without hesitation: no. The answer is simple because there is a previous question that we can also answer with a categorical no: has anyone ever seen a straight line before? Nobody can ever have seen one, since a straight line is infinite. At best, we can hope to see segments of a straight line, extremely long segments, indeed as long as we wish, but not infinite. In terms of parallel straight lines, the closest to this is the perspective acquired when we observe an extremely long stretch of a train track, for example. But even then, what we see, or what we think we are seeing, are two straight lines that meet at a distant point on the horizon. The fact that we appear to see this meeting point is the result of an optical illusion, since no matter how far along the track we go, we can never reach it. This is a commonplace experience, even one experienced by a high-speed train driver in his cabin when he points the train towards infinity at 300 kilometres an hour. In this respect, following the point at infinity makes about as much sense as trying to catch our own shadow.

What happens if, for example, instead of two parallel straight lines, we have three, or 10, or 20? For starters, we have what is referred to in geometry as a sheaf of lines, and most importantly, we have a direction. Imagine if we have a point at infinity on our plane – one of those points upon which two parallel straight lines converge. We can assign a direction on the plane to each of those points. In this case, all the points at infinity would represent different directions on the plane. In doing so, and also because they are so very far away, we can call this a line to infinity. This is a

somewhat pedestrian way of introducing ourselves to one of the most interesting and beautiful areas of mathematics: projective geometry.

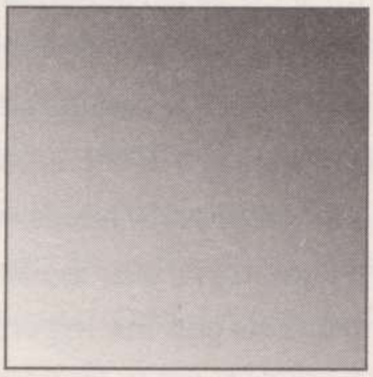
The basic idea lies in the fact that two parallel lines, or two parallel planes (which in affine geometry have the generic name *varieties*), do not share a common point. The only thing they share is their direction. This was realised by the geometers of the Renaissance, since they had already been working with representations in three-dimensional space for a long time.

The original idea of using a point at infinity came from Johannes Kepler (1571–1630), who was looking for a unified theory of conics. (He positioned the second focus of the parabola at infinity.) However it was developed more systematically by Gérard Desargues (1591–1661), who can be regarded as one of the fathers of projective geometry, an area that was not fully developed until the 19th century by the French mathematician Gaspard Monge (1746–1818).

Continuous transformations

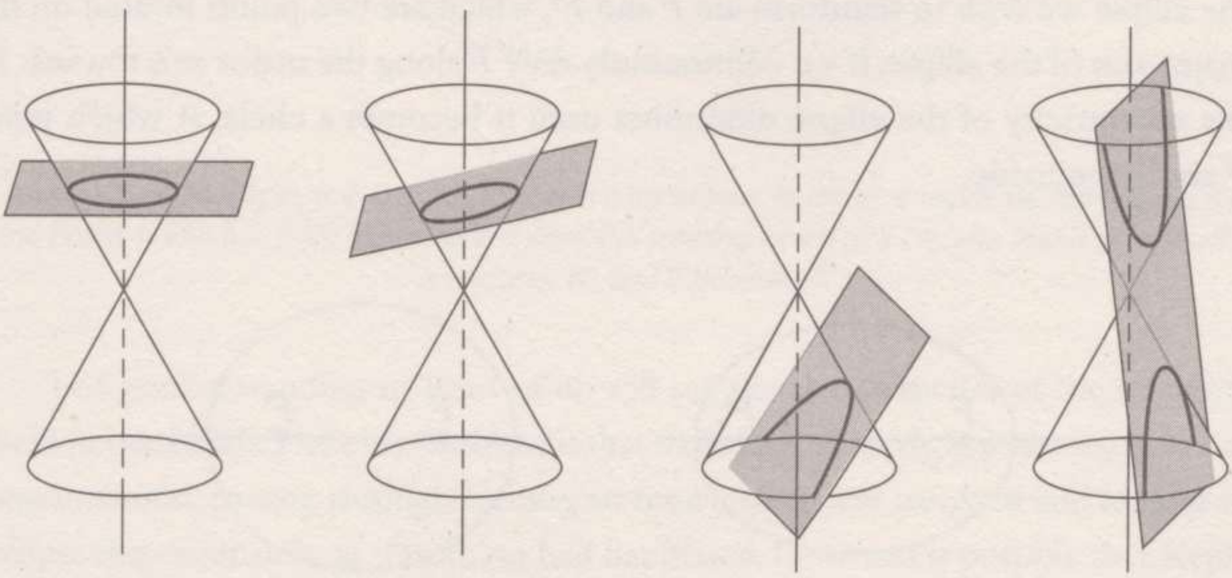
The concept of the infinitely divisible is inseparable from that of continuity. It is a complex subject. In the previous chapter, we considered the meaning of continuous, in contrast to discrete. Now let us consider the possibility of walking along the continuous continuously (that repetition is important). The most intuitive way of defining continuity is as follows: a line is continuous if we can draw it without lifting our pencil from the paper. In general terms, this term is also applied to the concept of a transformation. Imagine, for example, we had a parallelogram like the one below:

and we wished to use a continuous transformation to convert it into a square:



We must imagine that the sides are made of a malleable material, such as rubber, such that we can go from one shape to another, without jumps or breaks, and hence shift shape continuously.

In 1604, Kepler published a small work, *Astronomiæ Pars Optica (The Optical Part of Astronomy),* as a supplement to an astronomical work on the theory behind the manufacture of optical instruments – a very high technology at the time. Kepler studied conic sections and possible continuous transformations from one into another. Conics are flat geometric shapes obtained from the sections of a cone (or slices through it), as shown in the following drawing:



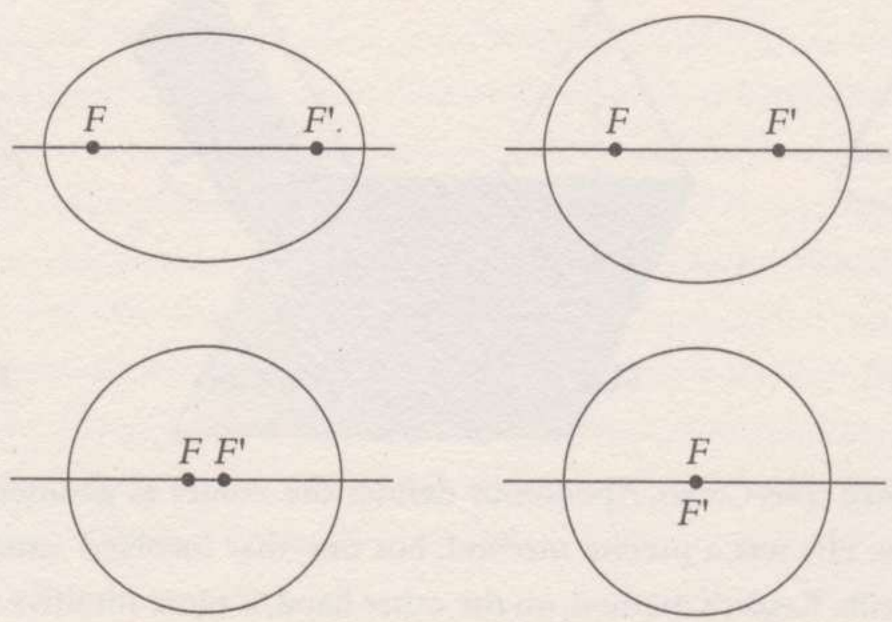
Circle Ellipse Parabola Hyperbola

In his work *The Conics*, Apollonius defines the conics as geometric locations on the plane. His was a precise method, but one that involved assuming certain geometric skills. Kepler's method, on the other hand, is more intuitive, and makes it easier to visualise the geometry. The definition is as follows: if we cut a cone (actually

two infinitely long cones facing opposite directions, such that they share the same axis and the points touch) with a plane perpendicular to the axis, we get a circle. By tilting the plane slightly, the circle becomes an ellipse, which grows in size as we increase the tilt of the plane. If we continue tilting the plane, we will reach a point at which it will be parallel to the generatrix of the cone, and the intersection of both sides will form curve known as a parabola. When, eventually, the plane is parallel to the axis of the cone, the intersection turns into the two branches of a hyperbola. All these curves – the ellipse, the parabola and the hyperbola – are referred to as conics (generally speaking, the circle is considered to be a special kind of ellipse). There are also other ways of cutting a cone with a plane to obtain the so-called degenerate conics (two straight lines).

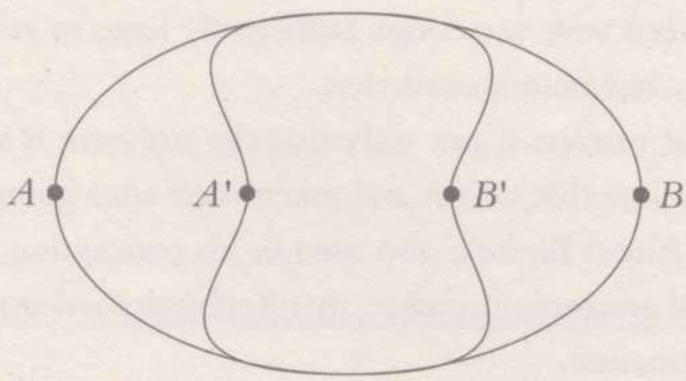
Imagine the movement of the plane that cuts the cone is being carried out continuously, such that there are no jumps. If we could visualise the transformation of the flat section, we would see how an ellipse turns into, for example, a circle or a hyperbola. Kepler established these transformations starting from an ellipse.

Remember that an ellipse is a conic that can be defined as the geometric location of the points of a plane whereby the sum of the distances to two fixed points, referred to as foci (one focus, two foci), is always the same. Imagine that the foci of the ellipse we wish to transform are *F* and *F'*, which are two points located on the major axis of the ellipse. If we continuously slide *F* along the major axis towards *F'*, the eccentricity of the ellipse diminishes until it becomes a circle, at which point *F* and *F'* coincide.



If, instead, we now move the focus F away from F', the eccentricity of the ellipse increases, which is the same as saying that the figure gets flatter (the eccentricity of an ellipse is a parameter e, between 0 and 1, which tells us precisely the extent to which the ellipse varies from a circle). A point will come at which the shape is no longer an ellipse, and becomes a parabola, a conic with a single focus defined by Apollonius as the geometric space made up of the points of the plane equidistant from one fixed point (the focus) and a fixed straight line referred to as the directrix.

If the long path from the point F does not stop at infinity and continues, it will wrap around the space and appear again to the left of F', giving us a hyperbola. Explained in a more visually intuitive way, to go from an ellipse to a hyperbola, we must pick up the ellipse at both ends, as if they were two handles, and bend them backward, as shown in the following drawing:



Starting from an ellipse, it is possible to obtain a hyperbola. To do so, imagine we are holding it by the points A *and* B *in both hands, as if it were the steering wheel of a car, and fold it in on itself so* A *becomes* A', *and* B *becomes* B'.

The person standing in front of us will see the two branches of the hyperbola held in our hands. The only problem is that to do this properly, it is necessary to turn space around, passing through infinity, to return to where we were and look at the ellipse dispassionately, as if nothing had happened. How was it possible that Kepler, a strong believer in a finite universe and opposed to any school of philosophical or mathematical thought that would defend actual infinity, was able to suggest such a transformation without even batting a metaphorical eyelid? Perhaps, speaking somewhat frivolously, we could say that Kepler was behaving like a 'turncoat' with respect to the theories of infinity at the time, in order to safeguard certain practical interests. It goes without saying that we are not referring to worldly, mercantile interests, but ones at the heart of applied mathematics.

The concept of continuous application that we have set out overleaf would become the cornerstone of projective geometry. The idea is as follows: Imagine we find a geometric property of the ellipse, whereby if we move one of its foci in the way explained above, the property is preserved. What else happens if the ellipse becomes flatter or longer? If the transformation is continuous, we will reach a point at which the same property can be applied to the circle, the parabola and the hyperbola.

The method of continuous change was subsequently used by Blaise Pascal (1623–1662), but applied to regular polygons, hence transforming, for example, a hexagon into a pentagon based on continuously moving two contiguous vertices until they came together.

How does Kepler get around the problem created by his method of taking a step through infinity? He simply makes the following argument: a line extends indefinitely at both ends until they meet at a common point. Hence, for Kepler, the universe was finite, but very, very large. Sufficiently large so as to hold everything we need and more... but finite nonetheless.

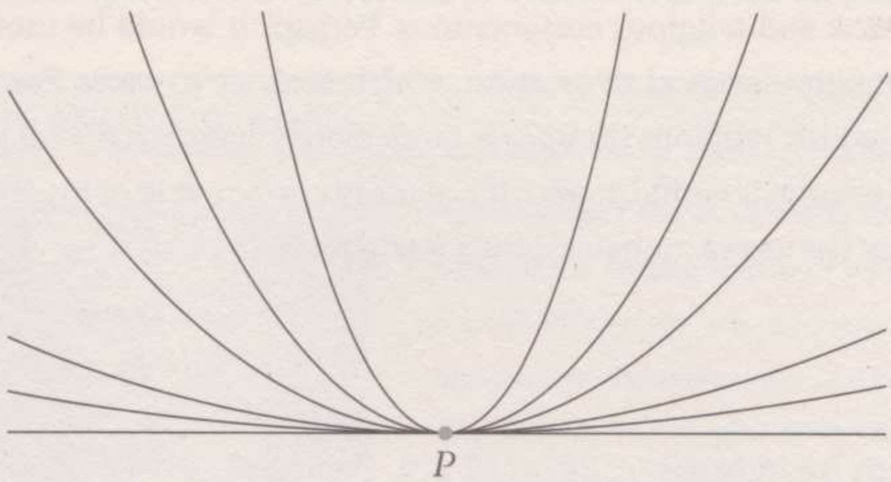
Regardless, what matters is not only that the universe is sufficiently large as to be able to hold a line that curves and meets itself after going round everything that exists (an idea Albert Einstein also used in his conception of space-time), but also, perhaps even of greater importance, that Kepler tiptoed around the concept of continuous transformation.

Quadratures

To carry out a quadrature means to transform a given shape into a square with the same area as the original shape. It is simple to calculate the area of a square, hence why we say that something 'squares' when it fits a given schema. One of the most common practical goals in applied mathematics has always been the calculation of areas. We figured out how to calculate the areas of flat figures bounded by straight line segments without too much problem. Pythagoras' theorem and Euclidean geometry made it possible to calculate the area of triangles and all kinds of quadrilaterals, and more complex shapes could be broken down into combinations of these two components. Of course, this required some skill and a lot of ingenuity, but in the majority of cases, the problem could nonetheless be solved. The matter became extremely complicated when one of the sides that bounded the shape was curved,

since suitable techniques had not been invented for such shapes. The Greeks had shown their skill at such calculations. However they were also accompanied by an uncomfortable journeyman whom they could not shake loose: actual infinity.

Why is it that when we stop dealing with straight lines, we can guarantee that infinity will rear its head, with all the problems that entails? Because when it comes to solving the problem, curves must be calculated as the end result of their approximation by an infinite number of straight line segments, or rather, a straight line that is the end product of a process of approximation using increasingly open curves, as shown here:



As the curves open up, the distance between them and the straight line grows increasingly smaller, especially in the areas surrounding the point P. *At infinity, the straight line and the curve are the same thing.*

Imagine a line with a point P that lies on it and a series of increasingly open (with less curvature) curves the tangents of which are P and which provide increasingly close approximations to the line. It is clear that no matter how many curves there are with a tangent at point P, none of them will be identical to the original line. We can imagine something like this happening and the infinite curves turning into a straight line. Potentially, this is possible, however *actually* (in the sense of actual infinity) we do not have a rigorous mechanism to guarantee this. Here, the need for infinity rears its head. The set of curves that increasingly approximate the line

have a shared property: they all have a quantity that is larger or smaller which we refer to as their curvature. In the limit step, when they turn into a straight line, this property disappears (though we could talk of zero curvature), and therein lies the change of nature. This is one of the reasons why infinity is associated with the mystery of creation. At some point in the space and time to which we do not have access, the same transformation occurs and one of the curves becomes a straight line. However we must bear in mind that this last phrase about one of the curves is both incoherent and inconsistent: there is no final curve, since in this case the concept of the infinitely small would disappear and we would make a discrete jump between the final curve and the straight line instead of making a continuous process. This act of creation has had considerable influence on scientific thought due to its philosophical and religious connotations. Perhaps it would be more reasonable to talk of 'mutation' instead of creation, which leads us to more Eastern forms of reasoning in which religious thought is more closely integrated with philosophical ideas. In this respect, it would appear, if not more correct, at least intellectually more elegant, to say the curve 'mutates' into a straight line.

Eudoxus

Together with Archimedes (c. 287–212 BC), Pythagoras (570–500 BC) and Euclid (c. 325–265 BC), Eudoxus (c. 408–355 BC) is one of the most important figures in Greek mathematics. In terms of conceptual mathematics he was doubtless the wisest of them all.

In those days, Greek mathematics was still raw from the setback represented by the discovery of incommensurable magnitudes caused by the appearance of irrational numbers. There were no clear criteria for comparing magnitudes of a different nature. It was Eudoxus who provided a clear definition for this purpose (as seen in Definition 5 of Book V of Euclid's *Elements*): "Magnitudes are said to be in the same ratio, the first to the second and the third to the fourth, when, if any equimultiples whatever be taken of the first and third, and any equimultiples whatever of the second and fourth, the former equimultiples alike exceed, are alike equal to, or alike fall short of, the latter equimultiples respectively taken in corresponding order."

Put more simply, this means that two ratios a/b and c/d are equal if, given any two natural numbers, k and k', the following hold:

$$\text{If } ka < k'b, \text{ then } kc < k'd;$$
$$\text{if } ka = k'b, \text{ then } kc = k'd;$$
$$\text{if } ka > k'b, \text{ then } kc > k'd.$$

While this may seem a trivial definition, that is certainly not the case. We must bear in mind, that, as stated by Eudoxus, there can also be ratios that contain the roots of numbers and even geometric shapes. For example, the first ratios could make reference to spheres, and the second to cubes constructed on their diameters.

EUDOXUS AND ASTROLOGY

Eudoxus was born around 408 BC in Cnidus, a city in the ancient region of Caria, in what is now Turkey. He was an astronomer and geographer and made important discoveries in both subjects – despite being poorly known compared to his classical contemporaries. He determined the trajectories of a number of heavenly bodies and determined that the solar year had six more hours than the 365 days it had been assigned to that point. Furthermore he was the first to divide the celestial sphere into degrees of longitude and latitude. He also drew a map of the sky and researched the creation of calendars, meteorology and seasonal changes in the banks of the Nile. His knowledge of astronomy led him to come into contact with priests who used it as a tool for astrological calculations. Eudoxus, who was strongly opposed to astrology, argued his case not based on a matter of beliefs, which were open to discussion, but on methodology: "When predictions are made regarding the life of a citizen using their horoscopes based on their date of birth, these should not be given any credit, since the influences of the heavenly bodies are of such complexity that there is no man on the face of the Earth capable of carrying out such calculations."

Furthermore, these rules contain the seed of what, in the 19th century, would become the definition of an irrational number, which Richard Dedekind established by what he called the 'cutting' method.

The other great contribution made by Eudoxus was the so-called 'continuity axiom', also known as *Archimedes' lemma* (Archimedes himself stated that it was actually Eudoxus who had revealed the lemma to him), which states: "Given two magnitudes having a ratio, one can find a multiple of either which will exceed the other." The great importance of this lemma lies in the fact that it makes it possible to use *reductio ad absurdum* to prove a proposition that is regarded as one of the most important in the history of mathematics, allowing Eudoxus and many other mathematicians who followed in his footsteps to calculate areas and volumes of curvilinear figures. Eudoxus' proposition is as follows:

"Two unequal magnitudes being set out, if from the greater there be subtracted a magnitude greater than its half, and from that which is left a magnitude greater than its half, and if this process be repeated continuously, there will be left some magnitude which will be less than the lesser magnitude set out."

This proposition also contains the essence of another great mathematical advance made during the 19th century, this time in the hands of Karl Weierstrass (1815–1897), which provides a precise and consistent definition of the limit.

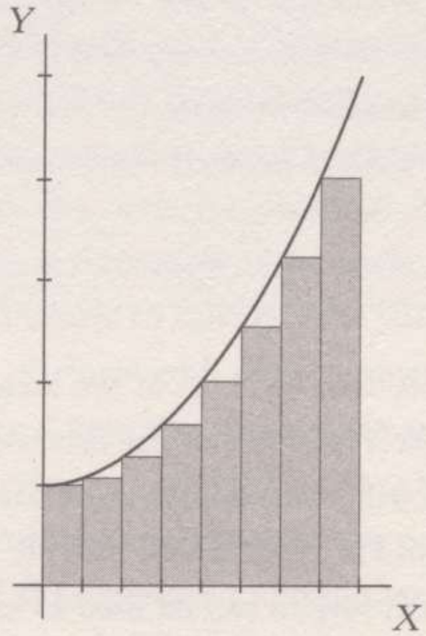
The Eudoxian method based on this proposition for calculating areas and volumes was known as the 'method of exhaustion'. It is not strange that many historians regard the era in which Plato founded his school of thought as a renaissance in Greek mathematics, since Eudoxus established the foundations that would support the eventual development of methods that give rise to what would finally become known as 'infinitesimal calculus'.

The method of exhaustion provided correct proofs if and only if the premises upon which they were based were solid (which they were in general). However, it had the drawback of not being a system that would produce new results. Remember that it was based on assuming a result to be true and analysing how it was possible to reach this result. It was known, for example, that the results for the volumes of the cone and the pyramid, for which Eudoxus reached satisfactory conclusions, had come from earlier mathematicians, such as Democritus.

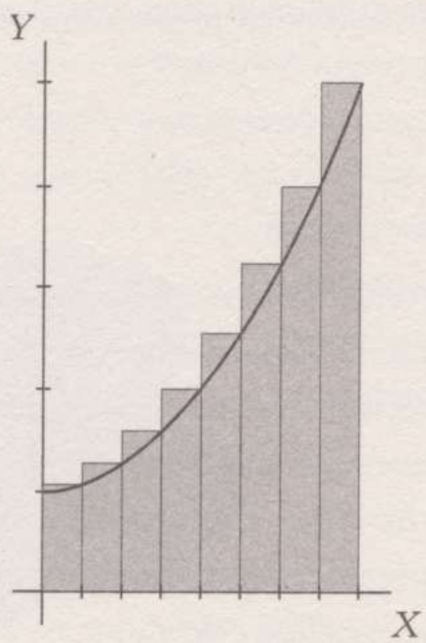
Today we have a method of integration that allows us to apply calculus using well-defined algorithms. That means that calculus can be done by a machine. The method is based on an idea that had been considered by Greek mathematicians. It is closely related to the previous example of approximating using rectangles. (To a

certain extent, the method of exhaustion used at that time corresponds to what we now know as the decomposition into Riemann sums.)

The method consists of establishing a series of rectangles with heights that do not exceed the height of the curve, i.e. rectangles with lower bases on the line and whose upper ends are beneath and just touching the curve:



The sum of the areas of all the rectangles arranged in this manner will obviously be less than the area we require. As the number of rectangles increases, their total area will grow increasingly closer to the area beneath the curve. The process can be repeated, only this time with the upper end above the curve:



INTEGRATION 'BY HAND'

There is a simple mechanical device, an 'integrator', that allows us to automatically calculate the area of a plane enclosed by a continuous curve. It is highly similar to those used to calculate distances on maps and consists of a small wheel with a tachometer that measures the distance covered as we follow the route of a path or a road. The integrator is very similar. When it covers the contour of a bounded surface, upon returning to the starting point, the device provides us with the value of the area enclosed within the curve. It is useful for mould designers, since it indicates the quantity of materials required to carry out their project.

This allows us to guarantee that the sum of the areas of the rectangles will be greater than the area we require. Now we can also increase the number of rectangles, and upon doing so, their area will better approximate the desired area, but this time the value will be slightly high. This system will provide approximations lower and higher than the curve. A similar system can be used to calculate volumes.

The results are compared with the value that is assumed to be correct (remember the method is based on analysing a previously obtained result) and stating an upper and lower bound. If these quantities are exceeded, we reach a contradiction. Later, in the 17th century this would become known as the *apagogic method*.

In both cases, the system inevitably leads us to consider the unavoidable presence of actual infinity, which in modern analysis would become the *limit step*, a step which, if applied by the Greeks to this and other problems of a similar nature, would have achieved spectacular results.

Kepler

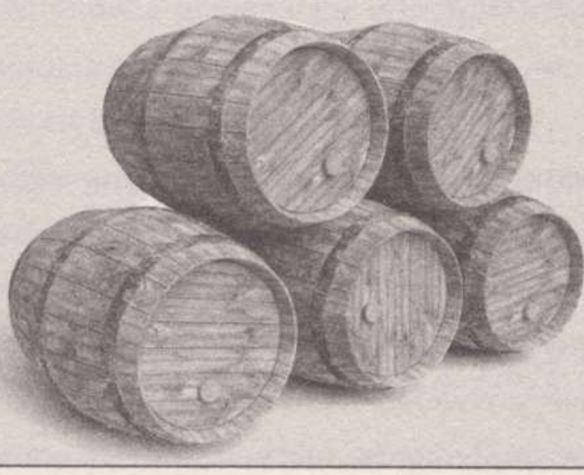
Kepler was one of the first Renaissance mathematicians to tackle the calculation of volumes. Moreover he did so in somewhat special circumstances, specifically on the occasion of his second marriage (his first wife having died the previous year) to Susanna Reuttinger. It was a marriage of convenience, since Kepler urgently needed a woman to take care of him, his children and the household tasks. Somebody must have warned Susanna of the peculiar nature of her future husband, because she was not surprised when he left the wedding celebration to fix all his attention on what a wine merchant was doing to the barrels that contained the guests' booze. The

receptacles were perfectly cylindrical and the remaining liquid inside was measured by introducing a stick obliquely through the lid. This made it possible for the wine man to deduce the volume of drink left in a barrel from the mark it left on the stick. The fruits of this observation was a work published in 1615 under the title *New Stereometry of Wine Barrels*. Kepler's solution to the problem was based on the method of indivisibles developed by Archimedes. It could be said that his wedding set the seeds of what would later become infinitesimal calculus.

However, it should be noted that Kepler's work in this area was more practical than theoretical, and in this respect it can be viewed as somewhat clumsy. For example, to calculate the area of a circle, it considered the sum of the areas of infinite triangles. Their vertices were in the centre of the circle and (infinitely narrow) bases lay on the circumference. Similarly, to calculate the volumes of spheres, it made use of the sum of the volumes of cones with their vertices in the centre of the sphere and their bases on its surface. Using this method, he arrived at the conclusion that the volume of a sphere was equal to one third of the radius multiplied by the surface. All these operations were justified for Kepler by the principle of continuity, which he had to take for granted if he wished to continue applying his method to the calculation of volumes.

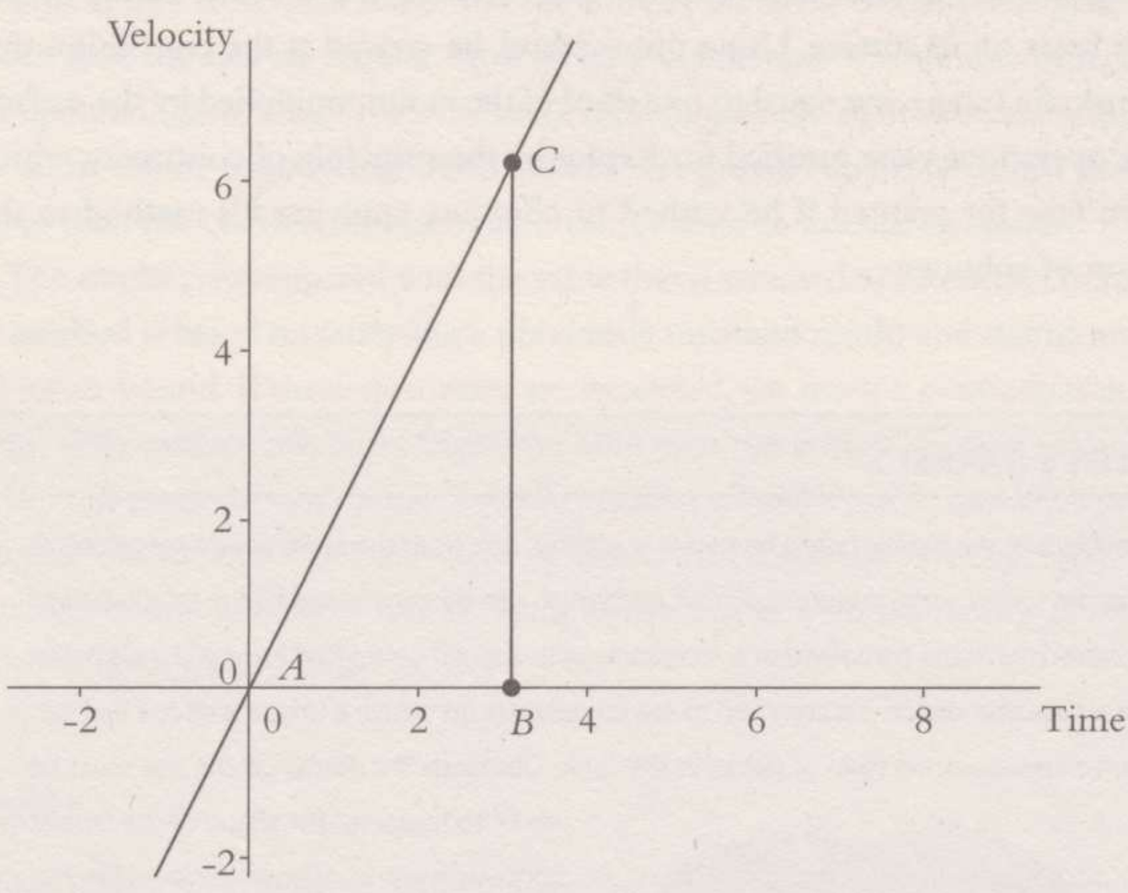
KEPLER'S BARRELS

The problem of the barrels posed by Kepler is a classic one from the collection of conundrums that can be solved using integral calculus, and which can be generalised to the calculation of the volume of a liquid contained in a receptacle with a given geometric form. When a petrol tanker arrives at a station it is common to see an operator introduce a large metal rod that can be used to measure the level of petrol in the tank. Obviously the marks on the rod must be made 'to measure' the shape of the tanker. Normally these containers have a cylinder-like body with ends that are finished with hemispheres or paraboloids.

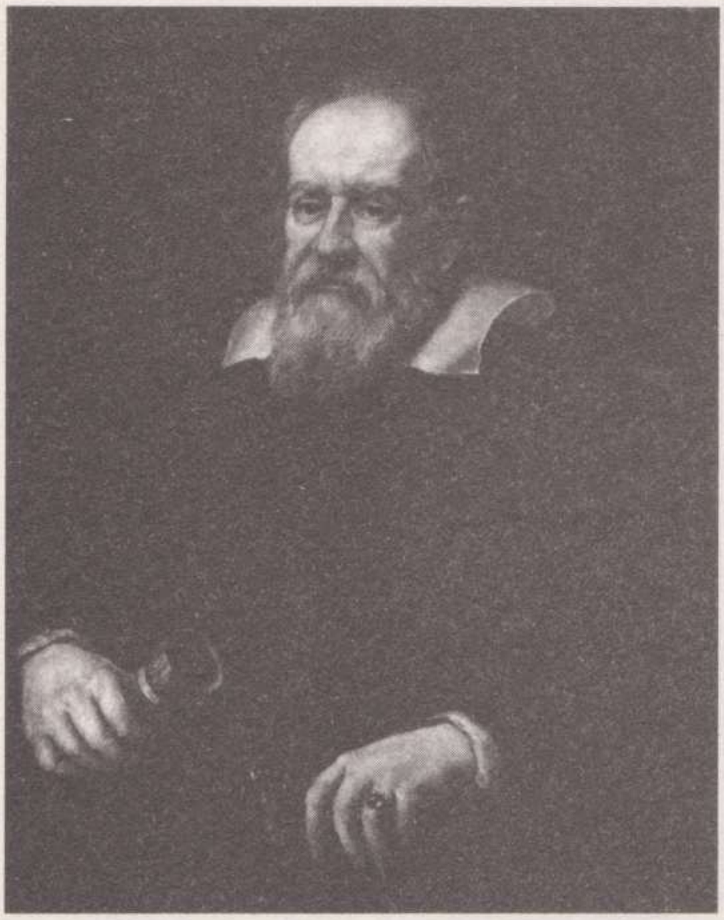


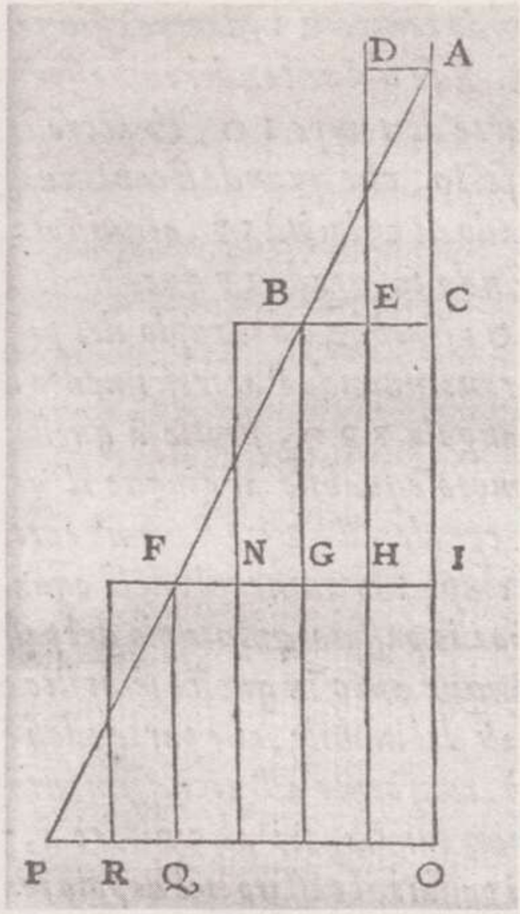
Galileo

Galileo Galilei (1564–1642) was a revolutionary scientist in many ways, and although the analysis of his extensive work or the influence this had on the history of science is beyond the scope of this book, we shall nonetheless give a cursory glance over his excursions into the realm of infinity. In first place, Galileo established motion as something that takes place without pauses, or rather advocated the continuous at the expense of the discrete. He knew this was a risky proposition, and that at some point he would have to accept the leap between potential and actual infinity. Moreover, in doing so, he succeeds in 'geometrising' problems of motion. Hence he considers that in a movement without a constant speed it is possible to obtain a geometric representation such as the following:



The horizontal axis represents time and the vertical axis speed. A non-uniform movement is of the form $v = 2t$, for example. This means that the speed increases as time passes: after one second it is 2, after two seconds it is 4, etc. Given a triangle ABC in which the segment AB represents the time that has elapsed and the side BC the speed, he concludes that the distance covered will be equal to the area of the triangle ABC. Galileo was interested in applying the method to more complex





A portrait of Galileo Galilei by the Flemish painter Justus Sustermans (1636) and a graph showing the motion of falling bodies.

motion, such as parabolic trajectories, which inevitably led him to come up against curves and the corresponding areas beneath them. His calculations used methods that were incredibly similar to those of Kepler. However, as we shall see further on, it would be Cavalieri, one of Galileo's pupils, who would be the first to establish a rational method for calculating this kind of surface.

As we mentioned, Galileo was bound to come up against the paradoxes of infinity, something that would necessarily lead him to reflect on its nature. This was how he came across a paradox for which he could find no solution. Formally speaking it was not actually a real paradox, but as we shall see further on, it indicated a possible mathematical definition of infinity.

This problem/paradox appeared in 1638 in Galileo's manuscripts and is set out as follows:

Begin with the set of natural numbers:

$$0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, \ldots$$

Now write the series formed by their perfect squares:

$$0, 1, 4, 9, 16, 25, 36, 49, 64, 81, 100, \ldots$$

It is clear that both sets are infinite, in the sense that it is possible to continue adding terms without reaching a limit. Furthermore, Galileo observed that for each element in the first set, there was another in the second set that was its square. Yet on the other hand, it seemed clear that there were many more numbers in the first set than in the second. Galileo then asked himself if the first infinite set was larger than the second, which led him to the apparent paradox. Either this is untrue, or, as Galileo himself reasoned, the arithmetic of comparisons based on the concept of greater, lesser and equal does not apply when we come to talk of infinity. And in this respect he was right, as Georg Cantor would confirm some three centuries later: "Infinity should obey a different arithmetic than finite numbers."

Cavalieri

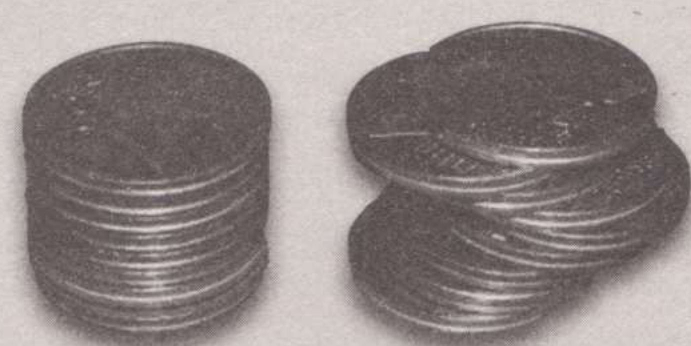
Bonaventura Cavalieri (1598–1647), a Jesuit mathematics teacher at a school in Bologna, was one of Galileo's disciples who took greatest interest in calculating areas and volumes. In 1635 he published a study on the problem entitled *Geometria indivisibilibus continuorum nova quadam ratione*. The title practically says it all. On the one hand, he accepts the principle of continuity and, on the other, he is willing to consider that continuous objects are limited to being divided into constituent parts, or 'monads,' like the atoms of matter can no longer be divided into smaller parts. To do so, he claims that a line is made up of points, like the beads of a necklace, and volumes made up of planes, in the same way that pages make up a book. This is to say that the indivisibles of a line are its points; those of a plane, its equidistantly spaced lines; and those of a solid, a set of parallel planes, all separated by the same distance. Cavalieri was aware that there had to be an infinite number of these indivisibles, although in reality he is another example of a mathematician who tiptoed around this issue. Moreover, he called his method the 'method of infinites,' but referred to his work as a treatise on indivisibles.

What is now known as Cavalieri's principle is stated as follows: If two bodies have the same height and their flat sections have the same area when taken at a same height, they also have the same volume.

Cavalieri used this method to prove that the volume of a cone is one third of the volume of the cylinder circumscribed around it. It goes without saying that his

CAVALIERI'S THEOREM

The method used by Cavalieri for calculating volumes can be visualised as follows: imagine we have two piles of coins, each of which has the same number of items. Let us now form two equal towers using each of the sets. Now deform the second tower, by sliding some of the coins on top of each other, such that it is no longer a cylinder. The problem of calculating the volume of the second tower, which is no longer a regular shape, would be rather complicated. However, Cavalieri's theorem assumes that both towers have the same volume. In this example, each of the coins represents an 'indivisible'.



According to Cavalieri's theorem, the volume of both piles of coins is the same.

method was widely criticised by his contemporaries, criticisms to which Cavalieri was unable to respond, because his method lacked a mathematically coherent justification. In his favour, it should also be stated that it did not claim to be rigorous, merely practical, and in this respect, he was successful. Mathematicians such as Fermat, Pascal and Roberval used it without undue scrutiny, the latter to great effect, leading to the discovery of the area enclosed beneath an arc of a cycloid.

Descartes

René Descartes (1596–1650) is the founder and key figure of rationalism. *Discourse on the Method*, was his most famous work, and the phrase "I think therefore I am" remains his best known, since it is the only truth with which he believed it is possible to begin his journey through systematic doubt. His method is, as its name sug-

gests, a set of rules that make it possible to argue correctly in any sphere of thought. There can be no doubting then that Descartes was more of a philosopher than a mathematician, and that his results in this field can be regarded as just one of the products of his broader method. (The fact that the sciences are currently separate from philosophy is not to say it does not exert any influence on them; it is just that today we are less aware of this influence.)

Aside from his other major achievements, such as the classification of curves for the identification of conics, it is in his *Geometry* where Descartes makes his most significant contributions to mathematics. Descartes believed that solving geometric problems often required excessive imaginative effort to recreate the continuous shapes in the mind's eye. This led him to create a system for conceiving of them as a set of points, each of which could be assigned a name or set of coordinates. In this way, a geometric problem could be converted into an algebraic problem, and many issues from algebra could also be solved using geometric methods. This work was the decisive step in developing analytical geometry and has been known as 'Cartesian geometry' ever since.

Descartes states the problem of the infinite in his work *Principles of Philosophy*. However, throughout the work – written in religiously fraught days – he does not refer to it as infinity, only the 'undefined'. Nevertheless, he makes no qualms about recognising the existence of the infinitely large, by stating that there is an undefined number of stars, and the infinitely small, by assuring us that matter is indefinitely divisible. This replacement of one term by another is not a coincidence, but is fully intentional, and Descartes justifies it by claiming that the word *infinity* should be solely reserved for making reference to God. He does, however, accept the possibility that things that are undefined can have a limit, just one that is inaccessible to us. In this way, Descartes reduces the impossibility of actual infinity to the limitations of human beings. That did not stop him from stating the existence of a potential infinity, since, according to his own argument, we cannot think of the finite if there is no infinite: "It could not be possible that my nature were as it is, i.e. finite but endowed with an idea of the infinite if the infinite being did not exist. The idea of God is like the mark of a craftsman imposed upon his work and it is not even necessary for this mark to differ from the work itself." Hence, for Descartes, the idea of the infinite is innate.

THE DANGERS OF PRIVATE CLASSES

Close-up from an 18th-century painting showing Descartes in the court of Queen Christina of Sweden, by the French painter Pierre Louis Dumesnil. The National Museum of Versailles.

In 1649 Queen Christina invited Descartes for a long stay in Sweden after much insistence that she wished to be taught philosophy by the master. Descartes saw an opportunity to escape from a political climate in the Netherlands in which the philosophical disagreements were beginning to turn violent. Legend has it that the queen had a penchant for cold rooms, often receiving her governors in rooms in which all the windows were open. (The chill kept official meetings short.) Descartes found himself required to deliver his classes under these intolerable conditions, aggravated by a timetable that was at odds with his habit of spending a lot of time in bed. The engagement proved disastrous: A carriage would pick Descartes up at half past four in the morning and take him to the palace to give classes to the queen half an hour later. Five months later, the great philosopher was struck down by pneumonia which killed him on 11 February 1650.

Chapter 4

Calculus

The history of mathematical analysis is one of the most fascinating in the history of mathematics, and its gradual development is certainly very closely related to conflicts with infinity, and more specifically, the mysteries of the infinitely small; it is not for nothing that it is also known as 'infinitesimal analysis'.

Infinitesimal analysis

Why *analysis* and why *infinitesimal*? 'Analysis' indicates a method of working that consists of tackling a problem by considering its solution as a working hypothesis then analysing the elements that have made it possible to reach the envisaged solution. Without a doubt, Descartes is one of the most famous figures when it comes to this method. Even its initial detractors would come to adopt the method which is rooted in so-called synthetic geometry, which dates back to the days of Euclid.

Secondly, we say 'infinitesimal' because of the concepts it involves. Essentially magnitudes associated with geometric elements are susceptible to being divided as many times as we wish (infinite division) and then used as indivisible elements and constituent parts of a whole. As we have already seen, this method began with Eudoxus and his famous method of exhaustion. It was given a more systematic treatment by the mathematicians of the 17th century, the key figures of which include Roberval, Barrow, Newton and Leibniz.

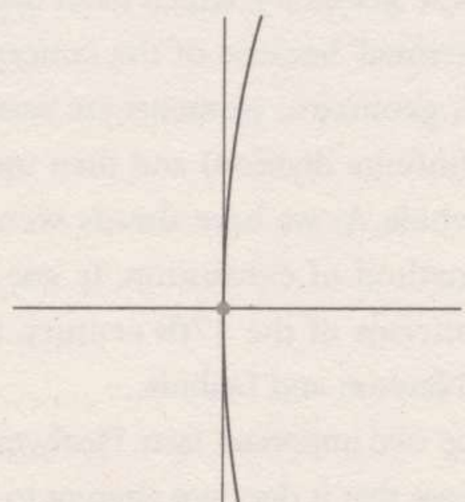
The subject also throws up two important facts. Firstly, mathematics becomes a self-sufficient discipline, in the sense that it does not attempt to adapt to models of nature. Nevertheless, the opposite is also true. Nature can be adapted to mathematics, not for the purposes of human vanity, but to enable solid theories that produce precise and practical results. For example, analytical methods can be used to determine whether the trajectory of a projectile is parabolic, since this is a clearly defined geometric shape (in analytical terms of a function). It is most likely that the projectile will not fit the path of the envisaged curve perfectly. However, to paraphrase Torricelli, that's the projectile's problem.

The second significant fact is felt in theoretical physics, insofar as it confirms the presence of two new concepts: 'body' and 'material point'. The first is attributed

to Descartes and the second to Newton. The apple that allegedly hit the latter on the head was not primarily a juicy fruit, but a 'body' with certain dimensions and inertia, or rather a mass, which, for the purposes of analysis, could be reduced to a 'material point'.

We should also bear in mind that at that time the development of physics was focused on applications. The problems discussed were defined by certain requirements of a practical nature. For example, in optics it was known that the angle of incidence and the angle of reflection were equal, something that was and still is essential in building optical instruments. However, angles are measured from a perpendicular reference line – the normal – that was positioned on a single point. This is no problem when the surface is a flat surface, since all we need to do is draw the normal as a line perpendicular to the surface at the given point.

If the surface is a curve, as is the case with the majority of optical instruments of interest – not least lenses – we are faced with a geometric problem of considerable complexity. As shown in the diagram below, the normal to a curved surface at a given point is a straight line perpendicular to the tangent of the curve at that point. At that time nobody knew how to construct the tangent of a curve at a given point.



The straight-line tangent 'touches' the curve at just a single point. The perpendicular to this straight line tangent at this point defines the normal *of the curve.*

Another example derives from the calculation of maxima and minima. Going back to the field of ballistics, the study of projectiles – especially cannonballs – the need to calculate the maximum range of a projectile as a function of the angle of a cannon is obvious (and even, on some occasions, the maximum height).

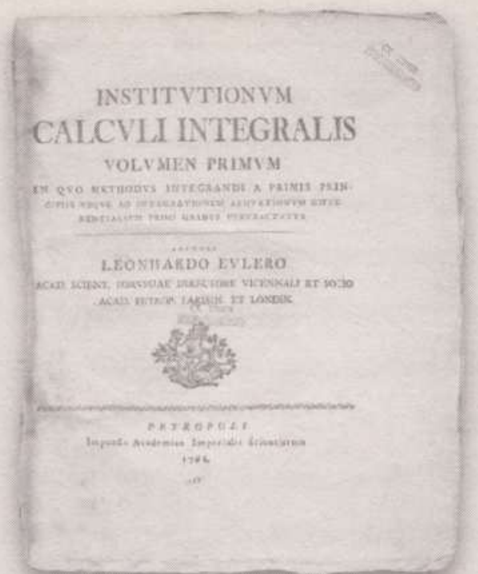
In summary, the four problems that needed to be solved and gave rise to the birth of infinitesimal calculus or analysis were as follows:

– The calculation of the tangent to a curve at a given point.
– The calculation of the maxima and minima of a function.
– The calculation of quadratures, being able to calculate the area enclosed by one or more curves.
– The straightening of curves, which consists of calculating the length of a curve between two of its points.

In all of these problems, infinity is present in its 'small' form, the infinitely small.

Newton and Leibniz are regarded as the scientists who realised how to combine their predecessors' knowledge to give form to mathematical analysis. Both followed

EULER'S INTEGRAL CALCULUS



Cover of the first volume of Euler's Integral Calculus.

It is possible to use integrals to calculate not only the area of flat shapes, but also the length of a curve, the volume enclosed beneath a surface, and volumes of bodies of revolution. Generally speaking, everything that can be described by an infinite sum of infinitesimal quantities. Integrals have so many practical applications that they constitute an entire branch of applied mathematics in their own right. Regardless of whether their effective calculation is carried out using small calculators or powerful computer programs, it is hard to imagine the work of an engineer without integral calculus in his or her mathematical toolbox. In 1770, the Swiss mathematician Leonhard Euler (1707–1783) gave a three-volume presentation of the subject. In some respects, modern calculus textbooks are merely modifications or stylistic revisions of this work, in which, 150 years after its publication, not a single oversight has been found to require correction. Move over Newton and Leibniz, Euler's *Integral Calculus* is the most important work ever written on calculus.

different paths, both had to confront the mysteries of the infinite, and both did so in different ways.

Newton

Isaac Newton (1643-1727), generally remembered more as a physicist than a mathematician, made extraordinarily important contributions to mathematical analysis. He devised an original system for dealing with matters related to quadratures and straightening curves. In doing so, he made use of infinite series, expressions defined by an equation the first term of which contains the function to be studied, and whose second contains an infinite sum of functions whose behaviour is known. For example, the first part of the following equation is a logarithmic function and the second a sum with an infinite number of terms expressed as power functions (parabolas) whose behaviour is known.

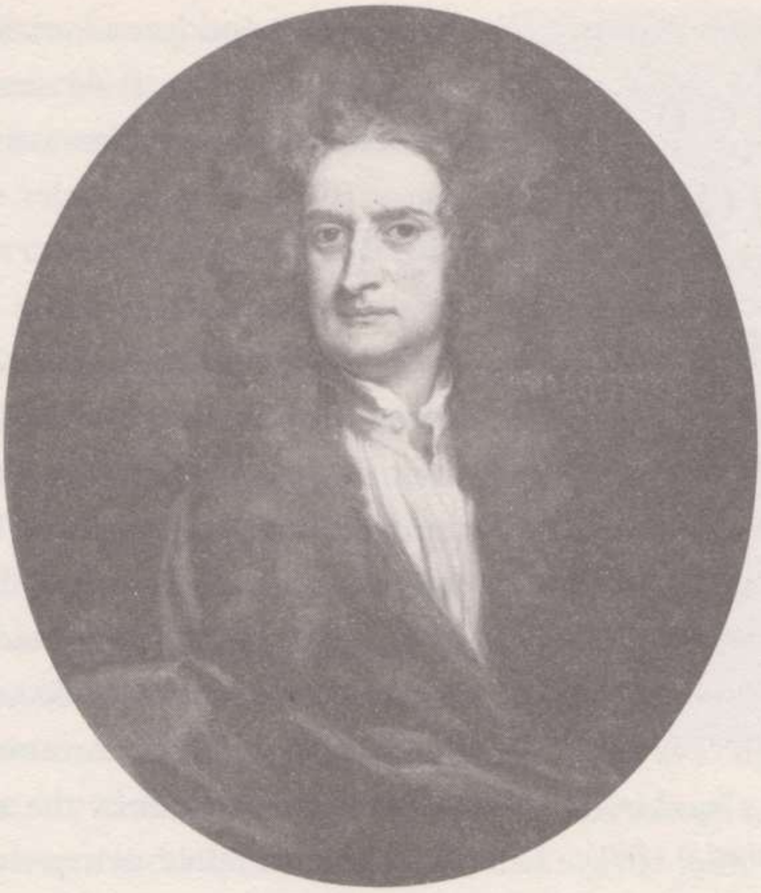
$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots$$

HIDDEN SCIENCE

Newton's *The Mathematical Principles of Natural Philosophy* has always been regarded as a difficult text, although this is not unusual, given that Newton was a difficult person who presented his work in intentionally difficult ways. On one occasion, Newton confessed to a friend that he had purposely made it tough "to avoid being attacked by small mathematical charlatans". Indeed, Newton had learnt his lesson from the incisive but not always fair criticisms he received on earlier works on the nature of light. He even came to write some of his results using cryptographic codes. The following series of letters and numbers

6a cc d ae 13eff7i 31 9n4o4q rr 4s 9t 12vx

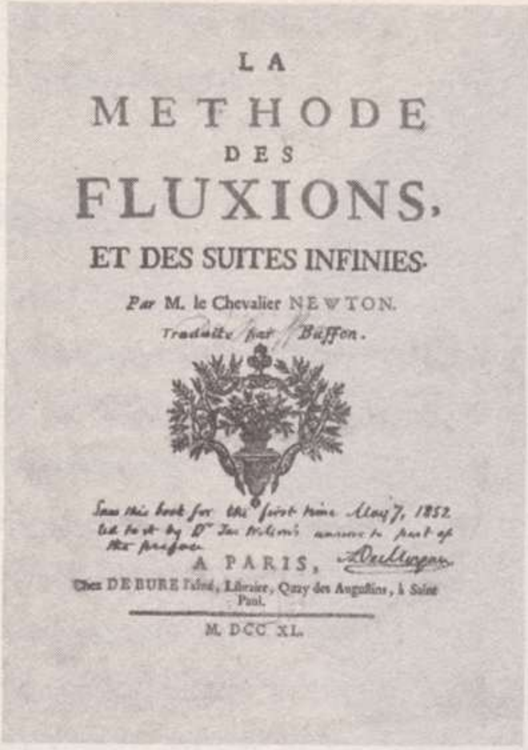
does not correspond to a complicated *password* or the serial number of a computer program. It is in fact a logogriph, a secret form of writing used by Newton to refer to his method of the calculus of fluxions and so prevent Leibniz being able to decipher it and claim authorship for himself. It has been is said that the latter would have required much greater wisdom to decipher the anagram than the secrets of infinitesimal calculus it bestowed.



A portrait of Isaac Newton by Godfrey Kneller.

Such expressions are referred to as infinite equations. The idea is that the more elements included in the second expression, the more precise the value of the function will be. If we only wish to carry out a simple calculation, it suffices to know the magnitude of the error. If we are trying to analyse a logarithmic function, classify it and study its behaviour, we must accept, albeit implicitly, actual infinity as the result of the sum of the series. Newton's only reference to this matter can be found in his work *De Analysi*: "For the Reasonings in this are no less certain than in the other; nor the Equations less exact; albeit we Mortals whose reasoning Powers are confined within narrow Limits, can neither express, nor so conceive all the Terms of these Equations, as to know exactly from thence the Quantities we want." Here we can once again see a pragmatic attitude. Actual infinity exists but it is beyond the limits of humanity. Nevertheless Newton can use this unknowable infinite term in his equations and accept the end results as valid.

However, it is in the second edition of his work *Methodus Fluxionum et Serierum Infinitorum* (1736, the first edition was published in 1672) that Newton makes use of the method known as *fluxions*. This represents an interesting development, whereby infinitely small elements are no longer considered static and are endowed with



LA
METHODE
DES
FLUXIONS,
ET DES SUITES INFINIES.
Par M. le Chevalier NEWTON.
Traduite par Buffon.

Saw this book for the first time May 7, 1852
led to it by Dr Jas Wilson's answer to part of
the preface.
A PARIS, A De Morgan
Chez DE BURE l'aîné, Libraire, Quay des Augustins, à Saint
Paul.
M. DCC XL.

A French edition of the Method of Fluxions *published in 1740.*

motion. It considers a variable as a point in continuous motion. It also imparts this quality to straight lines and planes and uses the term *fluent* to refer to variables that have these characteristics, and *fluxion* to refer to the result of this movement, or rather the comparison between two different states. We shall avoid going into the details of the method, but let us stress that Newton did not regard the use of infinitely small quantities to be necessary for his calculations, in light of all the contradictions this could entail. He considered these key quantities "…not as made up of small parts but as generated by a continuous movement. Lines are drawn, not by the addition of parts, but by the continuous movement of points…"

Using his method of fluxions, Newton was able to calculate tangents to curves, areas and lengths, and even maxima, minima and points of inflection for various curves. Moreover, he was able to do so while subtly evading the problems created by infinitely small quantities. However, for this he would have to pay a certain price. As established under these premises, analysis came to have significant limitations, and its future paved the way for other developments in which these strange infinitely small objects that were differentials were to lead the way. And all this in spite of the continuous presence of actual infinity.

Leibniz

The first mathematical works of Gottfried Leibniz (1646–1716) were on combinatorics, and although they bore the unmistakable hallmark of a genius, they were held back by the somewhat antiquated nature and mediaeval practices that still afflicted German universities at the time. In 1672, as part of an important diplomatic mission to which he had been entrusted, Leibniz moved to Paris. During the four years he spent in the city, he reinvented himself as a mathematician, partly thanks to the influence of Christiaan Huygens, who brought him up to date on the mathematics of the time.

Leibniz's first studies on infinite series date back to that time, with one of the most notable results being the series that bears his name, which establishes a surprising relationship between the number π and all the odd numbers:

$$\frac{\pi}{4}=1-\frac{1}{3}+\frac{1}{5}-\frac{1}{7}+\frac{1}{9}-\frac{1}{11}+...$$

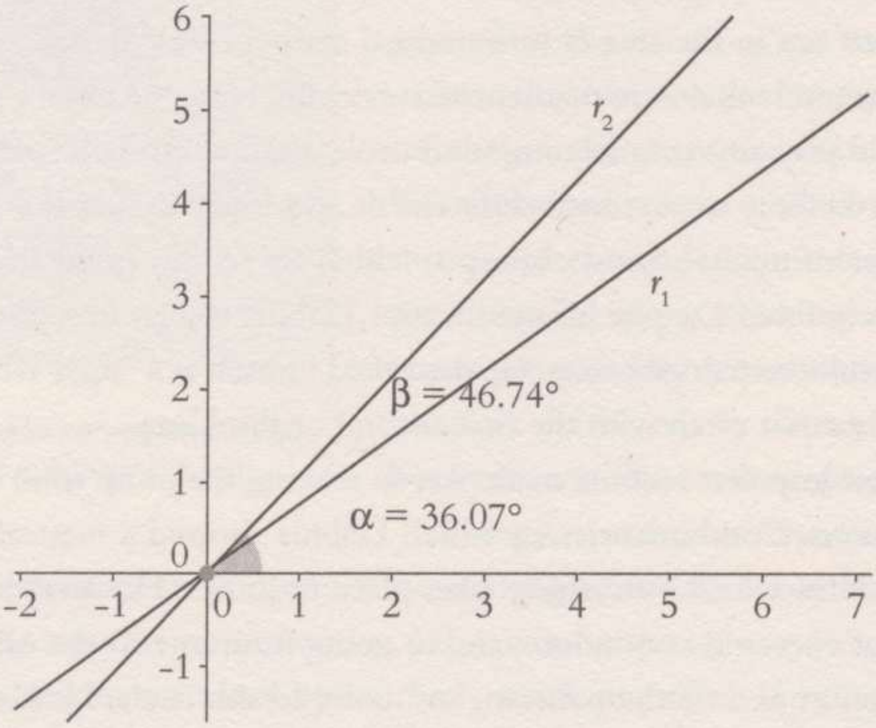
However, his most significant research was, without a shadow of doubt, the work he carried out in the area of infinitesimal mathematics, giving rise to one of the most important branches in mathematics: calculus. Here the choice of a suitable notation would play an extremely important role. Behind symbols such as d and $\int$, introduced by Leibniz to represent differentials and integrals, lies the synthesis of a large number of mathematical concepts, which up to this point had remained dispersed and confused. Despite his significance, Leibniz was far from thorough, and many of his results contained errors. He described himself as a "tiger who lets go of everything he cannot catch with the first, second or third leap".

The greatest leap that Leibniz made was in making the jump from the discrete to the continuous. Combinatorics, in which Leibniz showed a masterful ability, is a discrete world in which everything takes place in jumps. However the universe of functions, of curves, is continuous, and in going from one to the other, Leibniz showed his genius as a mathematician, and considerable audacity. He converted Cavalieri's indivisibles into new mathematical entities called infinitesimals, for which he then created specific algorithms. In a modern language and speaking in general terms, here we can see the key element upon which Leibniz based the nascent infinitesimal calculus.

AN APTITUDE FOR LANGUAGES

Leibniz, the son of a famous lawyer, was left orphaned at the age of six. He responded by becoming a precocious autodidact, especially in terms of his mastery of languages. Most of the books in the library he inherited from his father were written in Latin, a language to which the young Leibniz devoted all his efforts. At the age of 10, he was already reading the classics in Latin and Greek, and by the time he had reached 13, he was already able to compose hexameters in Latin in a record time. An aptitude for languages is an extremely common quality among famous mathematicians.

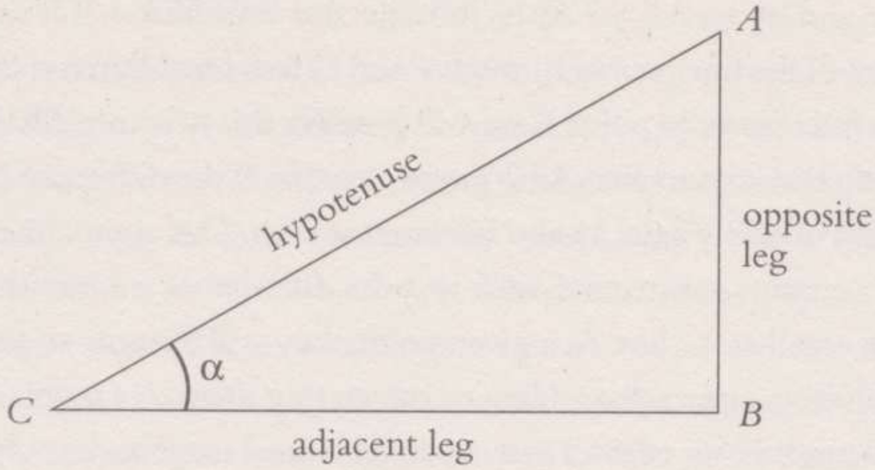
We know that a straight line is determined by two points, although it can also be determined by a point through which it passes and an angle. For example, the straight lines r_1 and r_2, which pass through the origin, are defined by the angles α and β, respectively. We refer to this angle as the *gradient*, a term which in everyday language is used to refer to a road or hill with a steep slope when the angle with which it meets the horizontal is large.



One way of measuring this angle is to make use of a protractor, which would give us a specific value, such as 24°. Another way is to calculate the tangent of the angle. For a right-angled triangle, such as ABC, the tangent of an angle can be obtained by dividing the length of the opposite leg by the adjacent one.

THE SEEDS OF INTERNATIONAL LAW

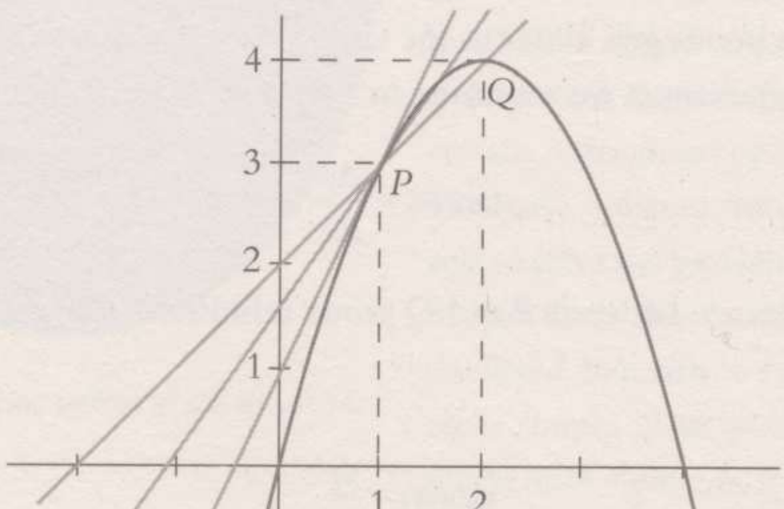
Leibniz began his studies in law at the University of Leipzig at the age of 15. Although he dedicated the majority of his time there to studying philosophy, by the age of 20 he was already able to win his doctorate in law, which was nevertheless denied to him by his faculty on account of his youth. He then moved to the University of Altdorf, where he completed his doctorate with a famous thesis on the historical nature of the law, a work that formed the basis of what would go on to become international law.



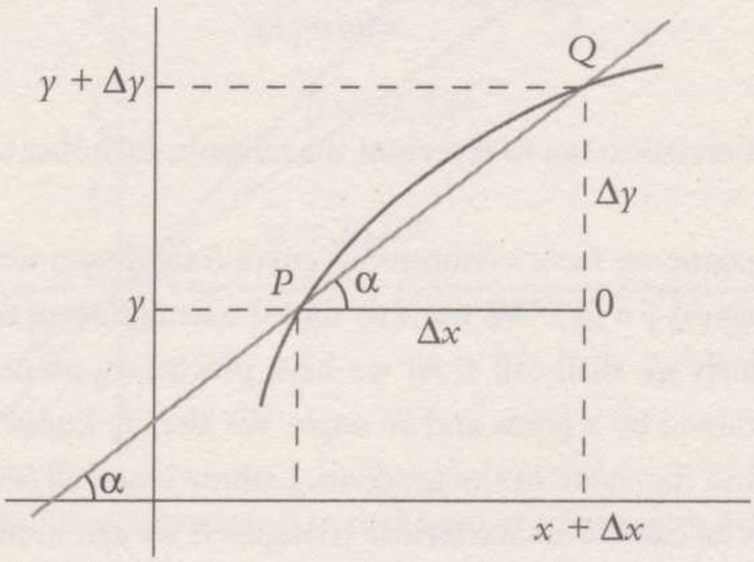
$$Tangent\,(\alpha) = \frac{opposite\ leg}{adjacent\ leg}$$

We use the abbreviation tan to represent the tangent, such that $\tan(\alpha) = AB/CB$.

Let us now imagine we have a continuous curve (one drawn without lifting our pencil from the paper) $y = f(x)$. We want to find the straight line tangent to this at a given point, which we shall call P. As we have previously mentioned, a straight line can be determined by a point and an angle. We already know the point, P. All we need now is find the value of the gradient. Leibniz based all his calculations on constructing what he called a 'characteristic triangle'. If we can stretch the metaphor enough, this triangle became the heart of the infinitesimal calculus.



Let us use x and y to represent the coordinates of the point P. Now take the point Q on the curve and let $x + \Delta x$, $y + \Delta y$ be its respective coordinates. It is easy to check that the gradient of the line passing through P and Q is defined by $(\alpha) = \Delta y / \Delta x$. If we now bring Q a little closer to point P, we will preserve this structure. All that changes is that everything becomes smaller. As Q grows closer to P, the difference Δx becomes smaller. Similarly, on the y axis, Δy also becomes smaller. This approximation can be carried out in a continuous manner, such that the differences we mentioned above can be made as small as we like. At a given point, they will become so small as to be insignificant values, meaning that adding or subtracting them to a number, regardless of its value will produce no effect. These are infinitesimal magnitudes, which Leibniz would later designate as dx and dy, respectively.



In the continuous process in which the point Q moves closer to the point P, the line joining both points gets closer to the tangent of the curve that passes through P, such that the gradient α we are trying to find is given by

$$\tan(\alpha) = \frac{\Delta y}{\Delta x}.$$

When the distance between P and Q grows infinitely small, the following will hold:

$$\tan(\alpha) = \frac{dy}{dx}.$$

LETTERS TO PRINCESSES

In many intellectual fields, Leibniz is better known as a philosopher than a mathematician. At the age of 20 he had already published his famous *Dissertatio de Arte Combinatoria*. In spite of the fact that many of his fundamental proofs can be found in publications such as *New Essays on the Human Understanding* (1703) or *Monadology* (1714), a large part of his philosophical thinking is in epistolary form, in letters addressed to princesses, specifically to three: Sophie, Sophie Charlotte and Caroline. Leibniz wrote his letters in a style, which, in addition to making clear his friendship, assumed a high level of intellectual preparation of the addressees, which appears to have been fully justified. The princesses were, to a certain extent, the only representatives of power with the possibility of creating scientific communities outside universities, in which the intellectuals of the time were stifled by the prevailing religious orthodoxy.

A portrait of Gottfried Leibniz at the age of 54.

This right-angled triangle, with legs dx and dy, is precisely what we referred to above as the 'characteristic triangle'. In fact, these infinitesimal lengths correspond to the straight sides of a polygon with an infinite number of sides into which the curve can be broken down. The big difference this time is that Leibniz handles these quantities as if they were numbers, albeit subject to certain restrictions, operating on them to obtain concrete results. He is even able to solve the problem of quadratures, in that he is able to calculate the area enclosed beneath a curve. To put it more simply, if the area is made up of differential elements, it suffices to add all these together to find the area we require (in this respect, differentiation and integration are inverse operations).

LEIBNIZ AND THE TEMPLE OF THE ROSY CROSS

At the age of 20, Leibniz was introduced to the mystic sect of the Temple of the Rosy Cross, something that should not come as a surprise (Newton and Descartes were also members). At that time it was hard for scientists to obtain all the information they required through official institutions. Experiments in alchemy were a basic requirement for joining this secret society, and Leibniz came to occupy the post of Secretary of the Brotherhood, assuming responsibility for duties such as transcribing the minutes of experiments and translating Basilius Valentinus' vast work on alchemy into Latin. Through the brotherhood, he came to meet Hennig Brand, the discoverer of phosphorus, helping him obtain the element from the urine of an entire regiment of soldiers for commercial exploitation. He also worked actively with Friedrich Hoffman, chair of Medicine at the University of Halle, in preparing the famous curative tincture referred to as Hoffman Drops, which can still be purchased in German chemists to this very day.

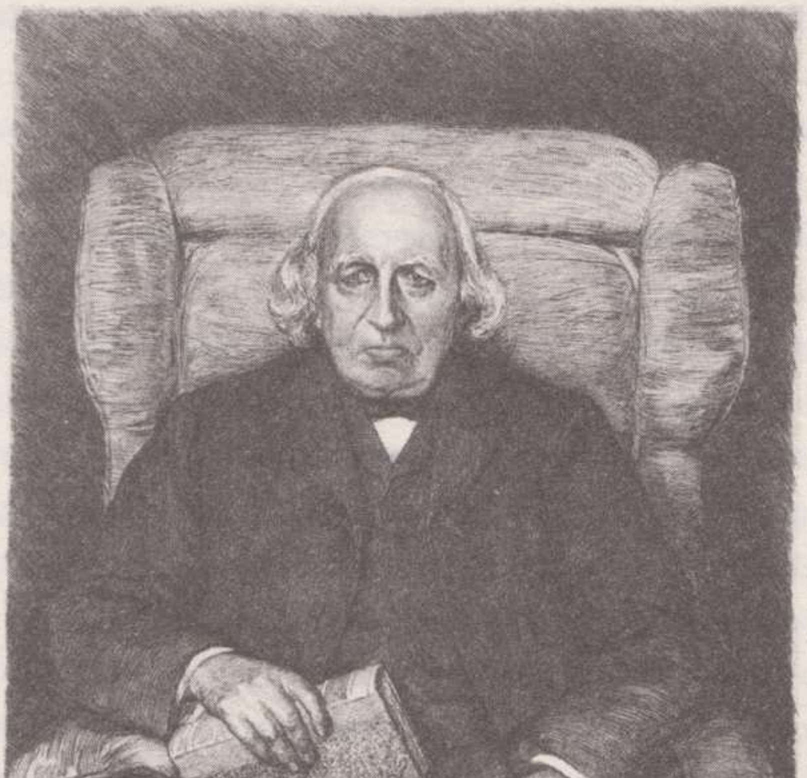
The Temple of the Rosy Cross, *by Teophilus Schweighardt Constantiens, 1618.*

The appearance of infinitesimals was not immediately accepted by the mathematical community at the time; indeed far from it. The characteristic triangle was there, but strictly speaking nobody could see it. It was an image, a representation of something that, once more, took place in the dark and impenetrable depths of the infinitely small, and that entailed, in spite of attempts to avoid it, the acceptance of actual infinity. Furthermore, it would be necessary to somehow skip over the Archimedean principle of comparing magnitudes, something that mathematicians such as Pascal, L'Hopital, Bernoulli and even Leibniz himself came to justify as highly special numeric quantities, which might vanish at a given moment. The title under which Leibniz chose to publish this work was telling: *On a Deeply Hidden Geometry and the Analysis of Indivisibles and Infinities.*

Epsilons

When we talk of 'epsilons' or the epsilon-delta method, we are not referring to the initials of a secret code or a plan of attack by the military but a complex mathematical apparatus directly related to the concept of limits. The concept was originally developed by Bernard Bolzano (1781–1848), but went unacknowledged, at least during his lifetime. The first to make practical use of the method was Augustin Louis Cauchy (1789–1857). However the mathematician who established it in its current form, in all its mathematical rigour, was Karl Weierstrass.

The definition of the limit using epsilons has a surgical precision in which it captures the infinitely divisible that for all intents and purposes seems impossible to overcome. For those who are untutored in mathematics, the definition is rather difficult. Nevertheless, this definition of the limit formed part of the secondary school curriculum for a long time. This is not to say that a student of secondary school age is not intellectually capable of understanding the complexities of epsilons, only it seems rather unnecessary – if not off-putting – to demand a full comprehension of the subject at that stage in their education. Most teachers skip over it nowadays.



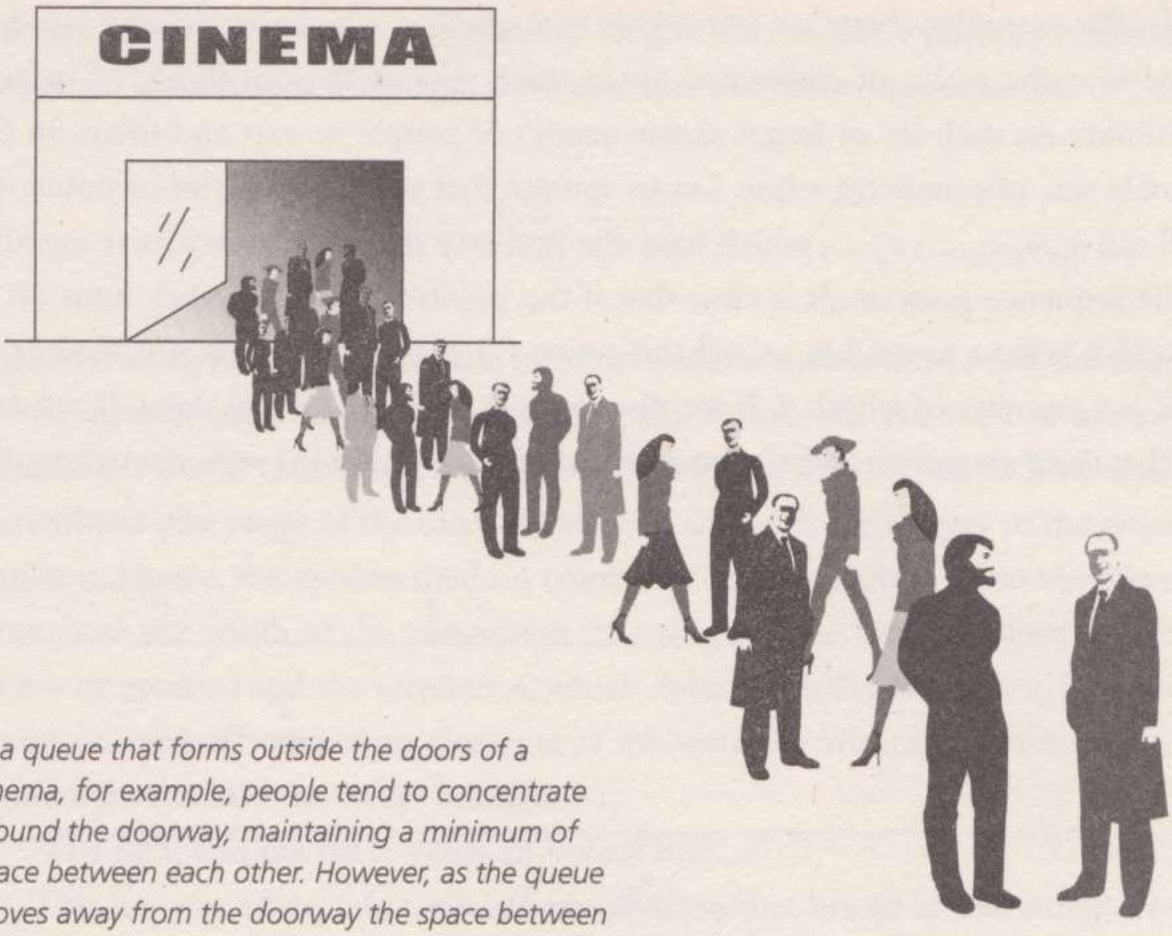
Karl Weierstrass in a print from 1895. The German mathematician was the first to make practical use of epsilons.

CLASHES OF GENIUS

As a source of setting and solving problems, letters are doubtless the oldest form of scientific communication in the world, and many have survived the centuries since their authors put pen to paper. Compared with other forms of recording ideas, a letter has the advantage of privacy. It is addressed to a single person or a group of people. Many scientific debates have been conducted by means of letters, and one of the most famous is the heated and lengthy conflict over the discovery of *calculus* by Newton and Leibniz. Newton arrived at his results fully independently and published his work prior to that of Leibniz. Newton was frequently selfish and unpleasant and could not stand his concept being shared or usurped by another. He accuse the German of plagiarism and thus gave rise to one of the most bitter disputes in the history of science.

We'll try to provide an intuitive approach to such a prickly subject. Essentially, the concept is closely related to the idea of accumulation. Imagine a queue of people starts to form in front of the door of a shop. We can see that the distance between the people and the door grows increasingly smaller, as does the distance between the people themselves. This is a natural phenomenon in queue formation: to begin with, when there are just a few people, the natural tendency is to maintain what we might call a comfortable distance between them. However, as the number of people increases, the distance between the people in the queue gets increasingly smaller. It is interesting to note that we are talking about two different, albeit closely related distances. One is the distance between the people and the door; the other is the distance between the people themselves. The latter increases as we move away from the queue. This is logical, since those who are arriving keep a natural distance, but as the queue moves forward, they feel pressurised by those behind them. All this can be summarised by saying that the people bunch together beside the doorway.

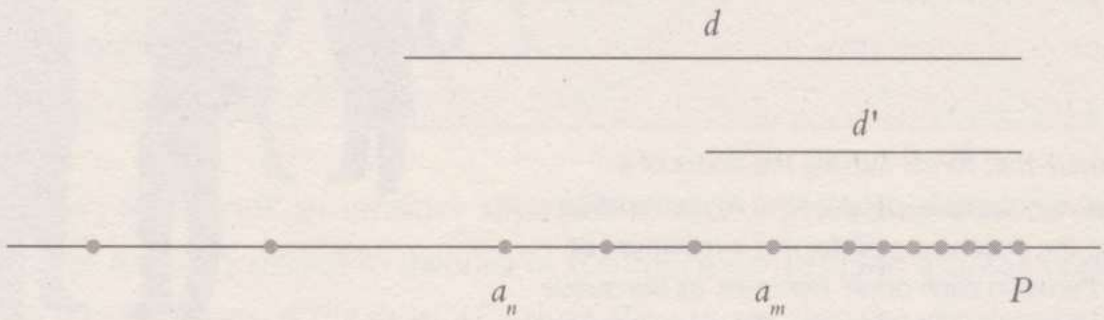
It is possible to establish the degree of accumulation by means of a parameter that measures, for example, the variation in the distance between consecutive people as we get closer to the queue. Under normal circumstances, this parameter will grow increasingly smaller.



In a queue that forms outside the doors of a cinema, for example, people tend to concentrate around the doorway, maintaining a minimum of space between each other. However, as the queue moves away from the doorway the space between the people increases.

By means of a 'measuring stick' it is possible to define the degree of accumulation, starting with a specific distance, for example 50 cm. The method for doing so is as follows. Place our measure at the door and if there are people 50 cm from the doorway, we obtain a certain degree of accumulation. We can use the length of our stick to confirm whether there are more or less people. We can also do this for the accumulation of people between them. Here we come across the first interesting point: the accumulation of individuals, commonly referred to as 'crowding', leads us to think of the presence of something causing this accumulation or, put another way, whether this accumulation takes place of its on accord, without a cause, or around a given object or situation. When we go out for a walk and see an accumulation of ants somewhere, we immediately think of the presence of food or the entrance to a nest. Another example is the accumulation of cars on a motorway. Are we approaching a toll booth or junction, or has there been an accident. This insistence on examples of 'accumulation' is justified by the fact that further on this will help us to understand one of the most interesting discoveries in the history of mathematics on the existence of certain numbers, numbers which, for decades, have remained stealthily hidden among the infinitely small.

In the examples above we have dealt with discrete sets, however now we shall enter into the realm of *continuums*, since these permit the possibility of infinite divisibility. As such let us forget about queues of people or cars and focus on the possible sets of points on a line. Let us assume that we have a series of points we shall call $a_1, a_2, a_3, \ldots, a_n, \ldots$, which have the property that they grow closer together as the sequence goes on. It is clear that if the points accumulate, they must do so around a certain point. Let us call this point P. Let us also assume our measuring stick is a segment of length d. If we place it with one end on the point P, we shall see that there are points of this sequence that remain within the segment of length d.



Furthermore, we can find a point a_n from which all subsequent points lie within the d. If we reduce the length of the segment to $d' < d$, we will now need to start from a point further on in the sequence, let us call this a_m, from which all other numbers will remain within the new bound. This is the epsilon method. It involves being able to ensure that for each quantity d there is always an n from which point all the elements in the sequence lie within the segment d. In this case, we can say that the sequence converges towards the point P. However, this is the same as saying two things: first of all, the series is infinite, and secondly the distance between the point P and the given term of the series can be made as small as we like.

When we are dealing with discrete sets, our discovery seems insignificant. For example, let us consider the sequence made up of the numbers 100, 50, 25, 12, 6, 3, 1 (we might imagine it as a queue of seven numbers and a doorway, which is zero). It is clear that the difference between any of these numbers and zero grows increasingly smaller, as does the difference between any two numbers. For example, there are 49 numbers between 100 and 50, whereas there are only two between 6 and 3. However, it is not possible to say that the terms of the sequence converge

on zero. Clearly, if we take a segment of length 1/2, we don't find any terms in the sequence around zero. However, if we consider the following sequence:

$$1, \frac{1}{2}, \frac{1}{4}, \frac{1}{16}, \ldots$$

there are always elements near the number zero, no matter how small the distance.

In mathematical terms, we are talking about intervals. An interval is a sort of parenthesis that is centred on the point *P*. The idea is that no matter how small the parenthesis (the range of the interval), we will always find elements of the sequence inside it. Hence, the epsilon method consists of playing with the two numbers that constitute the width of the parenthesis (the range of the interval, often referred to as ε – or epsilon) and the number n, which determines the element a_n from which point onwards all remaining elements of the series lie within the parenthesis. The game between both numbers becomes:

"For each epsilon, there exists an n such that…"

This is a way of dealing with infinite division that brings us extremely close to the mathematical definition of the concept of a limit. When we divided an interval in half an infinite number of times when discussing Zeno's paradoxes, we were establishing a numerical series like the one above. We can now apply the rigorous definition of the 'limit step', affirming that the last term in this series is the point zero. This is not to solve the paradox, it simply leaves it to one side, since we always come across the same: the points cluster around zero in an infinite series and we take it for granted that there is a final point, which is zero, despite the fact that in reality zero does not belong to this series. In reality, this jump is not justified, but it is well defined nonetheless. As Bertrand Russell would say, mathematicians are people who never know if what they are saying is true or not; normally they end up not knowing where they are going either, although they certainly have excellent knowledge of what they are doing.

In fact, Cauchy did not state the definition of a limit in the sense of elements that accumulate around a given point, but as elements that accumulate around themselves. That is to say, he did not have a tollbooth on the motorway, but a cluster of pile-ups here and there. Furthermore, things were not so simple, due to one important factor. If we only work with the rational numbers, the line upon

which we are placing the points is not full, it has a number of gaps. The situation is as follows: we have a series of points (we are now associating points on the line with rational numbers) such that it grows increasingly crowded. This is a phenomenon that can be mathematically defined and which was clearly established by Cauchy. The problem is that the crowding or accumulation may take place around a site or point on the line that is empty, or rather that does not correspond to a rational number. In this case, for example, the sequence

$$1, 1+\frac{1}{2}, 1+\frac{1}{2+\frac{1}{2}}, 1+\frac{1}{2+\frac{1}{2+\frac{1}{2}}}, 1+\frac{1}{2+\frac{1}{2+\frac{1}{2+\frac{1}{2}}}}, \ldots$$

which we defined in chapter 2, tends towards $\sqrt{2}$, which is not a rational number. In fact, using a suitably constructed right-angled triangle, we can ensure the hypotenuse falls on the appropriate spot, but this does not stop it from being a geometric construct. In Cauchy's time, it was about being able to define the points, or rather the numbers, arithmetically or analytically. The rational number had not been correctly defined, in fact it had not even been defined as a number at all. It would be necessary to wait for Dedekind and later on Cantor for an absolutely consistent definition. Of these two mathematicians, it was the latter who not only provided a correct definition but left the straight line without gaps. Earlier, we said it had a number of holes. In fact there were infinitely many, since, just like rational numbers, there is an infinite number of irrational numbers.

However Cantor merits his own chapter, since not only did he make the real line 'dense', but he came face to face with infinity in a way that nobody had been able to do before him.

Chapter 5

Cantor's Paradise

Perhaps it would be adventurous to say that there was one mathematics before Cantor and another after, but there are certainly those who make this claim. What is certain however, is that there was one infinity before Cantor and another after.

Fourier series

Jean-Baptiste Joseph Fourier (1768–1830) was a visionary mathematician and one of the pioneers who strode boldly into the new paradigm of mathematical analysis, developing one of the most fruitful theories in the history of applied mathematics. Among his works, one stands out in particular: *The Analytical Theory of Heat* (perhaps the most important of his publications), focussing on the problem of the propagation of heat, not only on account of its considerable scientific value, but also due to the fact it is regarded as the first ever work of mathematical physics.

The power series of a function consists of expressing it as an infinite sum of other functions. The importance lies in the fact that the functions that make up the sum may be easier to handle than the original function. Fourier series were not the first power series to be discovered, since by that time the Taylor series was also widely used, a series that required a function to be fully determined by its behaviour in a small interval.

Taylor series could be rooted in much more general functions, but suffered from the drawback that they had an excessively local character, in the sense that, once the behaviour of a series was known for a small interval, little or nothing could be said about its behaviour in any other disjunct interval. This led Fourier to study how to represent a function as the superimposition of other, simpler functions, generally speaking sinusoidal in nature, forming a new mathematical discipline known as *harmonic analysis*. The waves into which a function could be decomposed were referred to as harmonics, hence the name given to this type of analysis.

The possibility of breaking down a function into a sum of trigonometric sine and cosine functions represented an enormous mathematical advantage, since these functions were easily managed and could be easily represented, with straightforward

derivatives and integrals. Fourier showed that any periodical function $f(x)$, subject to certain restrictions, could be expressed as an infinite sum of trigonometric sine and cosine functions. The Fourier series raised two important problems that were difficult to solve and affected the foundations of analysis itself. They were questions that have been and continue to be fundamental to mathematics and relate to the so-called theorems of existence and uniqueness. The questions are, firstly, under what circumstances can it be assured that there is a sequence that really converges on a given function? And secondly, in the event that such a function exists, is it possible to guarantee its coefficients are unique?

In 1870 Cantor established a theorem that provided a criterion for the convergence of a Fourier series, and the following year, a second, which was an extension of the first, relative to the uniqueness of such a series. However, he ran into a problem that was difficult to overcome: the theorem was not general and had exceptions, values for which it did not hold. And there were not just a few of these points, but there were sets with an infinite number distributed continuously throughout those for which the theorem did hold. He had come up against irrational numbers, and this had led him to pose a question that went far beyond Fourier series and, to a certain extent, also beyond the concept of infinity, despite being closely related to it. Cantor began to seriously consider the relationship between the continuous and the discrete in the

Jean-Baptiste Joseph Fourier.

set of real numbers. On the one hand, there was a straight line on which, owing to purely geometric considerations, points were distributed continuously, whereas on the other hand arithmetic showed him a discrete distribution. There was something not quite right, and this was nothing less than the definition of real numbers itself or, more specifically, irrational numbers. (See the section 'Sets of numbers' in the appendix.)

Fundamental sequences

Cantor developed his theory on real numbers in two stages. In 1872, under the title 'On the Extension of a Theorem of the Theory of Trigonometric Series', he considered the problem of the existence of irrational numbers in a somewhat technical manner, but did not develop a complete theory. However, it would be many years more until he published in his *Foundations of a General Theory of Manifolds (Grundlagen)* in 1883, that the concept of a real number would adopt a consistent mathematical development. This was due to, according to Cantor himself, a profound philosophical reflection on the meaning of the concepts infinity and continuity. Familiar as he was with the work of Cauchy and Weierstrass, Cantor knew that among the set of rational numbers ($\mathbb{Q}$) there were infinite sequences of numbers that did not converge upon any rational number. These were the sequences established by Cauchy in which elements converged but without doing so on a specific rational number. In chapter 2, we already saw an infinite series that converged on $\sqrt{2}$, which is not a rational number. We also saw that such sequences of numbers are characterised by the fact that their elements can come to be as close to each other as we wish. Cantor called these series 'fundamental sequences' (they are now referred to as Cauchy sequences, although the original name is maintained in some texts).

Cantor surmised that fundamental sequences had to converge on an irrational number and he used this criterion for the definition of an irrational number. Continuing the analogy from the previous chapters, Cantor observed that there were accumulations of cars on motorways, and wagered that this was due to the presence of tollbooths. Putting it another way, the points around which a given series of numbers converged and those that failed to do so around a rational number (the divisions of the ruler we used for measuring) always do so around an irrational number, such as $\sqrt{2}$, $\sqrt{3}$, $\sqrt{5}$ or π itself. The problem was that it was necessary to give such numbers an identity, to define them mathematically.

There are certain properties that must be met by sets of numbers to ensure they form a coherent system or, put another way, that they really are useful and behave as we would expect them to for the most basic operations. The first of these guarantees is that they form a closed system for addition, subtraction, multiplication and division. We expect that adding together two integers will also give an integer. The second refers to ordering, and states that given any two numbers, it is possible to say unambiguously that they are equal or one is greater than the other. The third property, relative to density, is slightly more sophisticated, and not met by all sets of numbers. It states that between any two numbers, there is always another; the natural numbers and integers, as we have already seen, do not have this property. For example, there is no integer between 5 and 6. However, as we know, rational numbers do have the property of density. Cantor knew that the new set of rational numbers he was to define using fundamental sequences must have these three properties. However, he was unable to prove them rigorously. He was aware that the numbers he was defining were an extension of the rational numbers and he assumed, with immaculate logic, that these properties would naturally be passed on to irrational numbers. There was also another problem. Different fundamental sequences could give the same irrational number. These and other observations were overcome further down the line, with the concepts of equivalence relations and quotient sets, which is one of the current ways of defining these sets of numbers.

Now let us turn our attention to the fact that Cantor was unreservedly making use of the concept of actual infinity to define something as specific as a number, which, after all, was nothing more than the limit of an infinite series of numbers. In his first works, he did not even make use of the word *limit*. Furthermore, nor did he mention numbers, but instead numeric magnitudes. Indeed, Cantor was highly aware that he was entering muddy waters and that in order to tackle the problem of infinity and continuity, he would first have to arm himself with logical and mathematical tools that were not available to him, leaving him no option but to develop them himself.

By extending the system of rational numbers $\mathbb{Q}$ a new set was created $\mathbb{R}$, to which Cantor gave the name the *real numbers*. There are those who believe that this distinction was due to the existence of the *imaginary* numbers, which were already known at that time. However, there are sufficient grounds to believe that Cantor's

term was directed elsewhere. In *Grundlagen*, Cantor makes use of the word 'limit' and abandons the phrase 'numeric magnitudes' to refer to the new elements instead as real numbers. It is an important distinction, since it signifies that he is willing to accept actual infinity not just as mere speculation, but as a 'real' mathematical object, as real as a whole number or a fraction.

The real line

A straight line is an infinite set of aligned points. Cantor established the real line following the steps already set out in previous chapters, marking an origin and taking a length as a unit of measurement. The origin is situated on the number 0; to the right are the positive integers and to the left, the negative integers. We now add the rational numbers, i.e. fractions, with the positive numbers to the right and the negative ones to the left. Recall that the introduction of rational numbers to the line causes it to acquire a property it did not previously possess, that of density, which guarantees that between any two rational numbers, there is always another.

We have already seen how, in the world of Greek mathematics, the appearance of $\sqrt{2}$ resulted in a profound crisis. The root of the problem lay in the fact that this accursed number had a clear geometric construction using a right-angled triangle the legs of which had the same value, and for which the length of the hypotenuse was an irrational number that did not fit the set of points for the line on which the unit of measurement had been established. Hence, the length of the hypotenuse was valid as a magnitude, but did not exist as a number. In this respect, it could be claimed that the real line had an infinite number of holes, gaps to which no number corresponded, and as such it was not continuous.

In principle, with the introduction of irrational numbers, all the points on the line have a number assigned to them, whether rational or irrational, making it a dense line without holes. It now has every right to adopt the name the 'real line'.

Furthermore, the claim that the line, as a geometric entity, has been completely filled with numbers without leaving any holes, remains a risky one to make. Reflections on this matter led Cantor to grow more interested in the concept of the continuous than in the infinite, leading him to define a key concept, as a first alternative to infinity – enumerability.

The cardinal numbers

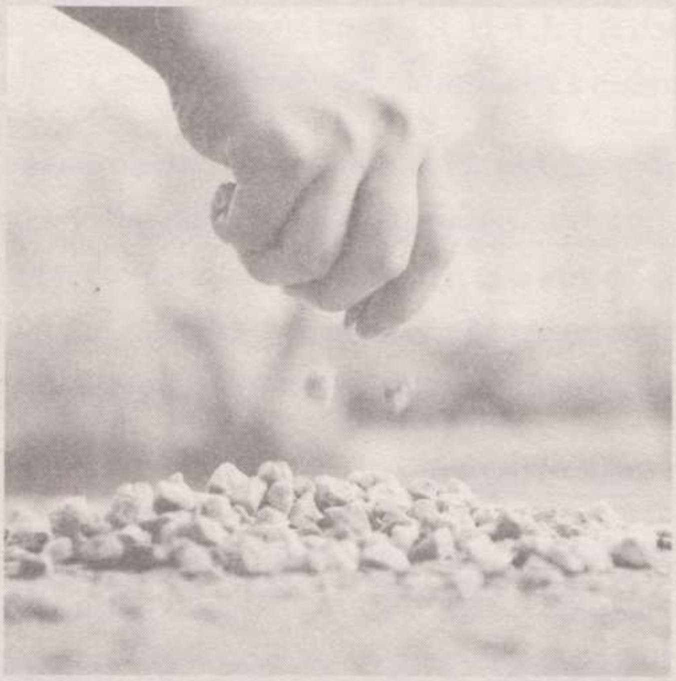
Cantor tackled the problem of being able to 'count' infinity. Until then, potential infinity had been defined by the possibility of adding new elements 'without a limit', but Cantor had proposed putting infinity firmly on the table, converting it into an 'act,' or actualising it. In other words using it as another mathematical element like any other. The simple exercise of counting a collection of objects, the most primitive arithmetic act, was to be completely revised and formalised, and two things were needed to achieve this: first, a correct definition of what we were talking about when we referred to a collection of objects, and second, a mathematical definition of what we called counting the objects in a collection.

Set theory would provide the solution to the first point: despite having already been sketched out by Bolzano, Cantor would develop a solid framework, making it possible to discuss the elements of a set in a wholly abstract manner.

Many historians of science have regarded Cantor's set theory as one of the most brilliant works of human thought. We shall not go into the details of the theory and all its complexities here, since for our purposes, it suffices to discuss a few highly intuitive concepts. However, we should note that the concept of a set is one of the most fundamental in mathematics, since all theory is practically founded on this concept. Henri Poincaré (1854–1912) once claimed that a mathematician was a person who dedicated their life to giving different things the same name. This is a succinct and slightly ironic way of expressing a great truth, since the fundamental goal pursued by mathematics is generalisation. And if there is one area in which this maxim can be fully applied, it is set theory, since the word *set* can designate anything that exists (and many things that do not). It is this generalisation that allows Cantor to begin to establish the rationale of actual infinity.

The first obstacle set theory comes up against is the definition of a set itself, since it is extremely difficult to discuss without using the word *set*, or a synonym: group, union, pile, etc. One of the best definitions, which does not use synonyms (at least apparently not) is that given by Bertrand Russell: "A set is a simultaneous consideration of entities." It is an interesting definition, because it suggests the concept as a mental attitude, a way of dealing with a primitive concept.

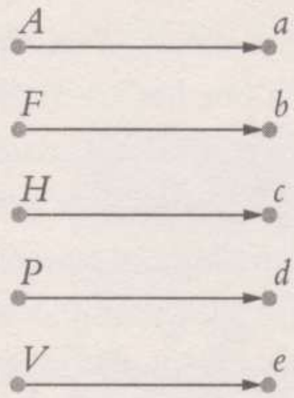
CALCULATION USING STONES



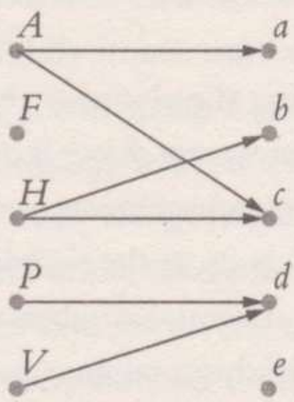
It is interesting to observe that historically, humankind learned to count before numbering systems came into existence. This confirms, in contrast to popular wisdom, that the concept of a bijective relation (see page 98) is at least as primitive, if not more so, than that of the number. For example, a shepherd who wished to keep track of the number of heads of sheep they were taking out to graze would need a bag full of small stones. For each sheep that left the pen, they would take a stone from the bag. As such, returning to the point, they would establish a biunivocal correspondence between sheep and a pile of stones that would allow them to tell if a sheep had escaped from the flock. (*Calculus* means 'little stone' in Latin, from which we then get the word 'calculate').

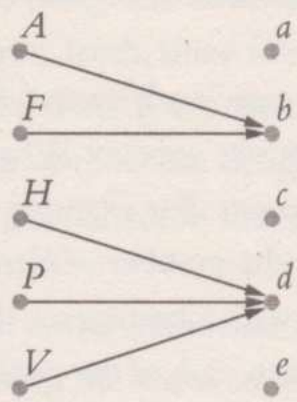
As we have previously noted, the act of counting the elements that make up a set is also a primitive concept. When we count, what we are doing is comparing the elements of two sets. For example, if we wish to know how many people there are in a given area (the number of elements of a set made up of the people in this area), we must begin with a known set, that of the series of natural numbers $1, 2, 3, \ldots$ and assign a number to each of the people in an ordered fashion, taking care not to repeat both numbers and people. When we are done, we look at the number that has been assigned, and if it is, for example 23, we say there are 23 people in the area. What we have done is compare two sets, that of the people and that of the numbers $\{1, 2, 3, \ldots 22, 23\}$, establishing what is referred to as a 'one-to-one' correspondence.

One-to-one correspondences can be established between sets of a different nature. However, what is important is that the rules of the game are preserved. For example, if we start with a set of capital letters, such as $\{A, F, H, P, V\}$ and another with lower case letters, such as $\{a, b, c, d, e\}$ we can establish a relationship between both sets as follows:



All elements of the first set must be related to one, and only one, element of the second set, in a reciprocal fashion. This is the only simple rule that governs this type of relationship, referred to in mathematical contexts as a 'bijection' or a 'biunivocal relation'. Correspondences such as the following do not comply with the established rule:





In this way, Cantor returns to the more primitive concept of the act of counting and establishes the concept of the cardinality of a set. If we observe the sets between which it is possible to establish a bijection, we can see that this is only possible if the sets have the same number of elements. For example, it suffices to attempt to establish an application of these characteristics between a set of four elements and another

with three, which is not possible without leaving unassigned elements or ones that have been assigned more than once.

Cantor then defines an equivalence between sets as follows: two sets have the same cardinality if and only if there is a bijection (one-to-one correspondence) between them." We say that both sets have the same cardinality or are *equipollent*, which is the same as saying they have the same number of elements.

Hence, if for a given set, such as a box of coloured pencils, which we refer to as A, we can establish a one-to-one correspondence with the set $N = \{1, 2, 3, 4, 5, 6\}$, we say that A and N have the same cardinality:

$$\text{Card}\ (A) = \text{Card}\ (N) = 6.$$

To some, it may seem that we are making extra work for ourselves in establishing what is, after all, obvious. However, looks can be deceiving, since this new logical weapon made it possible to provide a rigorous definition of an infinite set.

First of all, however, let us define a finite set. A non-empty set A (i.e. with at least one element) is said to be finite if, for a number n, the set A has the same cardinality as $\{1, 2, 3, \ldots, n\}$. Hence, the number n is precisely the number of elements in the set A. Otherwise, we say that the set A is infinite. Similarly:

A set A is said to be infinite if it has a proper subset B of A with the same cardinality as A; otherwise, A is finite.

The latter definition requires a more detailed explanation because herein lies the crux of the matter. First of all, what do we understand by a proper subset of a set? The idea is an extremely simple one. If we have a set A, such as $\{a, b, c, d\}$, a proper subset is any other set that can be formed using elements of A but without using all the elements, meaning that at least one must be absent. Examples of proper subsets of A are:

$$\{a\} \ \ \{a, b\} \ \ \{a, b, c\} \ \ \{a, c, d\} \ \ \{d\} \ \ \{b, c, d\}.$$

According to what we have said above, it seems logical we cannot establish a one-to-one correspondence between a set and a proper subset. The reason for this is also extremely simple. They do not have the same number of elements, since the subset will always have fewer elements.

Now let's turn that on its head and consider a case in which the above is actually possible. Imagine $\mathbb{N}$, the set of all natural numbers, and a proper subset, P, made up of

all the even numbers. It is clear we can establish a one-to-one relation between both sets, if we make each natural number n correspond to the number multiplied by two.

$$n \to 2n$$

Hence, we have,

$$\begin{gathered} 1 \to 2 \\ 2 \to 4 \\ 3 \to 6 \\ \cdots \end{gathered}$$

Or rather, for every natural number we have an even number, and conversely, for every even number, we have a natural number. This indicates that each set has the same cardinality. The claim that "there are as many natural numbers as even numbers" is not a paradox, despite presenting an apparent contradiction. This is an alternative definition of an infinite set: a set is infinite when it is possible to establish a one-to-one relationship between the set and part thereof (one of its own subsets)..

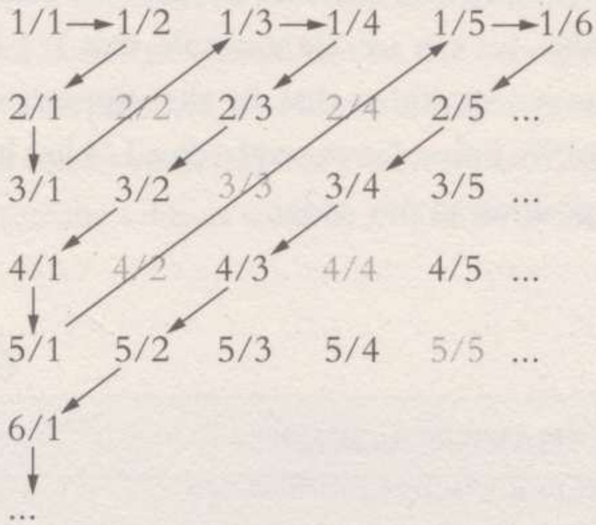
In this scenario, the 'paradox' suggested by Galileo (see Chapter 3) is no longer a paradox, but an affirmation that the set of natural numbers is indeed infinite.

We can use similar reasoning to show that the sets of natural numbers ($\mathbb{N}$) and integers ($\mathbb{Z}$) have the same cardinality. All we need to do is establish a one-to-one correspondence between both, making the positive numbers correspond to even numbers, and the negative ones to odd numbers. In this way, we can state that there are as many integers as there are natural numbers.

Enumerable sets

By means of this procedure, Cantor had established a new concept and one of great importance, that of the 'enumerability' of a set. By definition, a set A is enumerable if it is possible to establish a bijection between A and a subset of $\mathbb{N}$. Essentially, this is a simple idea that is used frequently in everyday life. When we say the tickets for a cinema screening are numbered, what we are really doing is establishing a biunivocal correspondence between a subset of the natural numbers and the seats, allocating a number for each seat. In a certain sense, we can say that the terms countable and enumerable are synonyms, since counting the elements of a set is after all a way of assigning a natural number to each element of the set in question.

It had already been shown that the set of integers was enumerable. The next step was to provide a highly surprising result: the set $\mathbb{Q}$ of rational numbers is also enumerable. Cantor shows this by proving there are as many rational numbers as natural numbers. To establish a correspondence between natural and rational numbers, Cantor makes use of an extraordinarily simple system. One can only assume that it did not occur to anybody before because they did not even imagine such a possibility, since it is goes against our most elementary intuitions.



The above system of diagonals devised by Cantor works as follows: construct a table of rational numbers (remember we are talking about fractions here). The first column contains fractions with a numerator of one. The second row has fractions with the numerator of two; the third row has fractions with a numerator is three, and so on. Once the table has been constructed, it is read in a particular way, starting with 1/1 and following the order indicated by the arrows. That route guarantees that we will read all the rational numbers once and once only. Hence, the one-to-one correspondence between natural numbers and rational numbers is established as follows:

$$1 \rightarrow 1/1$$
$$2 \rightarrow 1/2$$
$$3 \rightarrow 2/1$$
$$4 \rightarrow 3/1$$
$$5 \rightarrow 1/3$$
$$\ldots$$

The most amazing thing about this result is that it has established a one-to-one correspondence between a discrete set of numbers (the natural numbers) and a denser set of numbers (the rational numbers). Here infinity begins to reveal its fascinating mysteries. In principle, it is reasonable, or perhaps better put intuitive, to think all discrete sets will be enumerable. The surprising discovery lay in seeing that a dense set such as $\mathbb{Q}$ was also enumerable. Intuitively we associate the idea of enumeration with the possibility of finding the 'next' element for a given element, something that is impossible in a dense set, since none of the elements have a 'next'. Observing the previous table, we can see, for example, that 1/1 is the first number and 1/2 is the next. However, we know that, by the property of density, infinite numbers lie between them. We know, for example, that 1/4 lies between 1 and 1/2, and corresponds to sixth position in our order.

THINKING IS MORE THAN SPEAKING

According to Cantor's set theory, the set of words we can generate, either by speaking or writing, is enumerable. If we bear in mind that the set of signs in a given language is finite (letters, punctuation, etc.), it is clear it will generate an enumerable set. Contrast this with the set of things we can think, which is clearly not enumerable. We can imagine, for example, the set of circumferences on the plane with the power of the continuum. Extrapolating from this idea, what we can say can be ordered, whereas what we think cannot; at least not in its entirety. Hence, we must admit that part of our thought permits an ordering, although most thoughts occur chaotically.

abcdefghijklm
nopqrstuvwxyz

The letters of the alphabet represent a limited and hence enumerable set.

This fact meant that from Cantor onwards, the concept of enumerability became relevant to the concept of continuity. The next question that arose was inevitable: when we extend the set of rational numbers to include irrational numbers, do we still have an enumerable set? Can we say that $\mathbb{R}$ is an enumerable set

The answer is no. Let us now consider Cantor's proof for this. This time he made use of a method similar to that of the diagonals used to prove the enumerability of $\mathbb{Q}$, but considerably more complex. It was used in conjunction with *reductio ad absurdum* to show that the interval (0,1) of all real numbers between 0 and 1 was not enumerable, and hence nor could this be the case for $\mathbb{R}$.

Cantor's use of this method would set a major precedent that would play a decisive role in the mathematics of the 20th century. Without going into greater detail, it forms part of the schema used by Gödel to prove his famous theorem.

Beyond infinity

All know you, none span you; with this trick the moderate appears to be great, the great infinite and the infinite more.
The Hero. Baltasar Gracián (1601–1658)

Cantor already knew that neither the real line nor any of its segments were enumerable. His next step was an enormous one and it was there that he came face to face with infinity.

Let us recall that we obtain the real numbers by adding the set of irrational numbers to the set of rationals, the former being of type $\sqrt{2}$ or π (any number that cannot be obtained as the quotient of two integers). This set is also infinite and dense. However, it is not an enumerable set, such as the two previous ones, or rather it is not possible to establish a correspondence between the set and the series of natural numbers 1, 2, 3, 4, 5,...

Cantor then asked the following question: we have infinite sets with the same cardinality, which are equipollent, or rather they have the same number of elements, such as the natural numbers, even numbers and the rational numbers. However, in this scenario, we are also faced with the presence of a new set, that of the real numbers, which is also infinite, but which appears to have more elements than the preceding three sets.

This led Cantor to consider one of the most revolutionary ideas in the history of mathematics: are all infinities equal, or are there larger and smaller infinities?

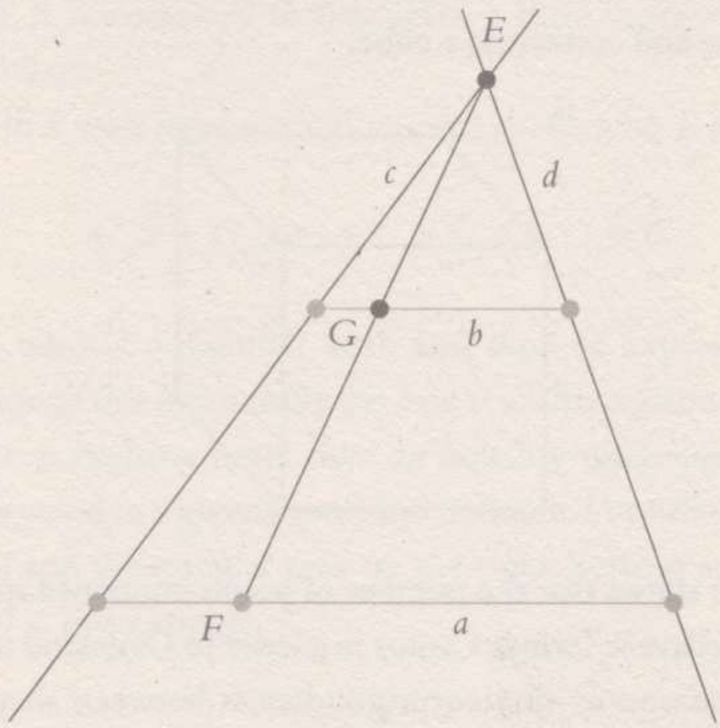
He already had one infinity to use as a starting point – the natural numbers. He then showed that the set of real numbers, $\mathbb{R}$, is not enumerable and contains more elements than $\mathbb{N}$. It is larger than the sets of natural and rational numbers. He decided to use the term aleph-one, symbolised by $\aleph_1$, to represent the cardinality of $\mathbb{R}$. Here was the birth of transfinite mathematics.

Cantor already knew that $\aleph_1$ is the number of points on a given segment of any line. Hence, regardless of their size, two segments have the same number of points. This may seem surprising, however the proof is extremely simple and, in fact, was already known by the Greeks.

A 9TH-CENTURY VISIONARY

Thabit ibn Qurra (c. 836–901) was a prestigious Arabian scientist. It is known that he came from Harran, a region of Anatolia. In addition to a large number of texts on theology and philosophy, he also produced an interesting work on mathematics, essentially dealing with arithmetic. In it he suggested the possibility that there were different types of infinity, in the sense of their ordering, that is to say that some could be larger than others, thus exhibiting an unusual audacity for the time. In this respect, his can be considered an authentic forerunner to Cantor's theories.

Given two segments, *a* and *b*, to establish a one-to-one relation between each of their points, one must do the following: join the two ends of both segments using two straight lines, *c* and *d*, which will intersect at a given point *E*.

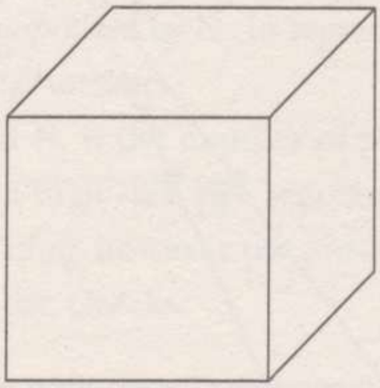


Now, given an arbitrary point F on the segment a, we can join this point with E, the intersection of lines c and d using another line segment. Point G, where it cuts the segment b is the image we are seeking. It is clear that in this way, for each point on the segment a we can obtain a point on segment b and vice versa. This shows that the number of points contained on both segments is the same.

It is here that Cantor performs a somersault. He constructs a square from one of these elements:

From here, he successfully shows that the number of points contained in the square is also a set with a cardinality of $\aleph_1$, hence the number of points contained

in the square is the same as the number on any of its sides. In his next step, he uses the square as a base and constructs a cube:



Once again, he shows that the number of points contained in the cube is $\aleph_1$. "I see it but I don't believe it," wrote Cantor in a letter to Dedekind in 1877, explaining the results of these one-to-one correspondences between shapes with different dimensions. Indeed, what Cantor had shown went against all intuition and even mathematical ideas on the concept of dimension itself, since it was dealing with the infinite points of objects with one, two and three dimensions, all of which had the same cardinality $\aleph_1$.

This result is more than surprising. What it is saying is that on a given segment, regardless of how small it is (the reader can bring their index finger and thumb together as when indicating something is extremely small), there are as many points as there are in the known universe. Inside the infinitely small lies the infinitely large.

In fact, we can go further still. $\aleph_1$ is the cardinality that corresponds to the number of points in any space and hyperspace. Put another way, for science fiction fans, in the event that we had any type of access to higher dimension spaces (four, five or any number of dimensions for that matter), $\aleph_1$ would continue to represent the number of points it contains.

Transcendental numbers

We have seen that the sets $\mathbb{N}$ (natural numbers), $\mathbb{Z}$ (integers) and $\mathbb{Q}$ (rational numbers) have the same number of elements (equipollent). This is an infinite number that Cantor represented as $\aleph_0$. The set of real numbers is obtained by extending the set of rational numbers to include the irrational ones. The burning question now is: are there enough irrational numbers so that when we add them to the rational numbers, the total jumps in such a way as to reach $\aleph_1$? The answer to this question

holds a mathematical curiosity that is not without a certain mystery. However, to understand it, it is necessary to first learn a little more about the so-called *transcendental* numbers.

An equation in x with degree n and rational coefficients is an equality when

$$C_n x^n + C_{n-1} x^{n-1} + \ldots + C_1 x + C_0 = 0.$$

To someone who is unfamiliar with this type of expression, it may seem complicated, although this is not really the case at all. An equation, as we understand it in this context, is nothing more than an equality made up on the left of the unknown value x raised to a given power and multiplied by other numbers (referred to as coefficients) and the number zero on the right. Solving an equation involves finding a value of x. For example,

$$x - 2 = 0$$

is an equation with coefficients 1 (assumed and therefore not shown) and −2, and whose solution is $x = 2$.

An irrational number, such as $\sqrt{2}$, is the result of solving an equation of the following type

$$x^2 - 2 = 0.$$

By definition, a number x is said to be *algebraic* when it is the root (solution) of a polynomial equation with integer coefficients. Let us clarify a few things to make this definition easier to understand. A polynomial equation is nothing more than a polynomial that is equal to zero, such as, for example

$$3x^2 + 5x - 1 = 0,$$

where 3, 5 and −1 are the coefficients. The expression

$$\sqrt{3}x^5 - 5x^2 = 0$$

is also an equation, however the first coefficient is not an integer and, as such, it cannot be regarded as a polynomial equation using the definition we have provided.

In contrast, the number three is an algebraic number, since it is the solution to the equation:

$$x - 3 = 0.$$

It is clear that any rational number is also an algebraic number, since it is always possible to construct a polynomial equation whose solution is that number. As we have seen before, $\sqrt{2}$ is the solution to the equation $x^2 - 2 = 0$ and hence is also an algebraic number.

When a number is not algebraic, it is said to be *transcendental*, a term coined by Euler to reflect the fact that the calculation of the number 'transcends' the normal operations in some way. Proving a number is transcendental can be an extremely arduous task. The French mathematician Joseph Liouville (1809-1882) showed the existence of transcendental numbers and found a method for generating specific cases of these numbers. The first to have the honour of appearing in this short list was L (the Liouville number), whose definition lies beyond the scope of this book. However, the number in question is as follows:

$$L = 0.110001000000000000000001000\ldots$$

In 1873, the French mathematician Charles Hermite (1822-1901), a student of Liouville, proved that e (the base of natural logarithms, whose approximate value is 2.7182818284 5904523536 0287471352...) was not an algebraic number. The proof that this is no mean feat is that Euler himself had failed in his attempts.

One of the most famous and popular numbers in the history of mathematics is π (pi), which expresses the relationship between the length of a diameter and a circumference. Proving that e was transcendental required such effort from Hermite that he lacked the motivation to attempt the same task for the number π, as he explained in a letter written to Carl Wilhelm Borchardt (1817-1880): "I have not dared to attempt the proof of the transcendence of π. If others try, none will be happier than I, however believe me, dear friend, the result will require considerable effort."

Hence, the mathematical community would have to wait a little longer for the proof that π was transcendental. It was provided by Lindemann in 1882 and represented a milestone in the history of mathematics. It also proved the impossibility of squaring the circle.

It has been shown that e, π, e^{π}, $2^{\sqrt{2}}$, $\sin(1)$, $\ln 2$, $\ln 3/\ln 2$, and a few other numbers are all transcendental. However there are currently important unresolved problems, such as discovering if numbers of the form e^e, π^{π} and π^e are transcendental. It is known, for example, that at least one of the two numbers (probably both) π^e and

Charles Hermite in a photograph taken around 1887. The French mathematician proved that the number e is not algebraic.

π^{+e} is transcendental, however it has not been possible to prove this for either of them. As can be imagined, transcendental numbers have proven strange beasts, and ones that are difficult to find, leading us to the conclusion that they must be rare. However, the reality is quite different: there are many, indeed, an infinite number.

In the infinite set of real numbers, we have, on the one hand, the rational numbers, which are all algebraic and, on the other, irrational numbers. The irrationals may be algebraic or may be transcendental, by definition there is no formula to tell one from the other. In fact, it turns out that the latter set is in the majority: there are more transcendental numbers than algebraic ones.

Displaying a surprising level of ingenuity (he himself would come to be surprised at his own results), Cantor provided an extremely simple proof that there was an infinite number of transcendental numbers. On the one hand, he knew that the set of real numbers is not enumerable, whereas on the other, he had proven that the set of algebraic numbers was enumerable. From both propositions we can deduce

immediately the existence of non-algebraic numbers. Cantor also showed that this set was not enumerable.

The conclusion is what turns the set of real numbers into a monster is precisely the presence of these elusive transcendental numbers.

Transfinite numbers

The arithmetic of transfinite numbers differs from that of finite numbers.
G. Cantor

As we saw in the previous section, if we have a set such as $A = \{a, b, c, d\}$, it is possible to form a series of subsets

$$\{a\}, \{b\}, \{c\}, \{d\}, \{a, b\}, \{a, c\}, \{a, d\}, \{b, c\}, \{b, d\}, \{c, d\}, \{a, b, c\}, \{a, b, d\}, \{a, c, d\}, \{b, c, d\},$$

which are what we call *proper* subsets of A. They receive this name due to the fact that when we speak of the subsets of a set, that is of sets formed using its own elements, the full set $\{a, b, c, d\}$ and the empty set are also subsets of A. The empty set is symbolised by $\varnothing$, the set with no elements, and is considered as a subset of any set. These two sets (the original set with all its elements and the empty set) are called *improper* subsets. Hence, when we add these two sets to the previous collection, we have the complete collection of all subsets of A:

$$\{\varnothing\}, \{a\}, \{b\}, \{c\}, \{d\}, \{a, b\}, \{a, c\}, \{a, d\}, \{b, c\}, \{b, d\}, \{c, d\}, \{a, b, c\}, \{a, b, d\}, \{a, c, d\}, \{b, c, d\}, \{a, b, c, d\},$$

which gives a total of 16 subsets.

We have $2^4 = 16$, such that the number of subsets of A is 2 to the power of the number of elements in A. It can be shown that this is always the case, such that we can state that given a set with n elements, the number of its subsets is always 2^n.

The set made up of all subsets of a set A is referred to as the *power set of A*, and is written using the symbol $\wp(A)$. Cantor proved that, in general, for any given set its power set was greater (i.e. it contained more elements) than the original set. This is the same as saying it has a higher cardinality. To avoid excessive use of parenthesis,

let us use an alternative symbol for cardinality: vertical bars, such that, from now on, Card $(A) = |A|$. Hence, we can express the above result as follows:

$$|A| < |\wp(A)|.$$

Cantor proved a number of theorems, but when we talk of 'Cantor's theorem', we are normally referring to this theorem in particular.

An alternative way of expressing it would be:

$$|A| < 2^{|\wp(A)|}.$$

This theorem makes it possible to 'order' infinities. Cantor believed the 'smallest infinity' is that which corresponds to the cardinality of $\mathbb{N}$, the set of natural numbers, which he called $\aleph_0$, such that, using the naming convention we described above, we have:

$$|\mathbb{N}| = \aleph_0.$$

Applying Cantor's theorem:

$$|\aleph_0| < |\wp(\aleph_0)| < |\wp(\wp(\aleph_0))| < \ldots$$

The series of cardinalities appearing in this ordered sequence was referred to by Cantor as aleph numbers, with a number being assigned for each: aleph-1, aleph-2, aleph-3, etc. They are pronounced as aleph one, aleph two, etc., and are written with their ordinal number as a subscript to the Hebrew letter aleph:

$$\aleph_0, \aleph_1, \aleph_2, \aleph_3, \ldots$$

These are the so-called *transfinite* numbers.

NEARLY INFINITY

Not just infinities or transfinite numbers go beyond our finite nature. For example, a number such as

$$10^{10^{10^{10^{10^{10^{10^{10}}}}}}}$$

is monstrously large. It could be the answer to a mathematical calculation. Using a suitable language, a computer processor could arrive at this result after a reasonable number of steps. This is due to the fact that we have the required tools both in terms of mathematical notation and programming languages. If this were not the case, if we had to write the number with all its digits, we would require a material (regardless of the media, whether it was paper or any other) considerably in excess of the number of particles in the universe. Furthermore, neither would we have the time required to write this number, since it is also far in excess of the age of the universe.

Any number that can exist, even those which we haven't even imagined yet, is contained in this ordered sequence of numbers. If, prior to Cantor, we had stated that nothing could be greater than infinity, after his theories we can be sure that there will always be an infinity larger than another. Cantor had surpassed the limits of creation. As vast as a god was able to create, there was always a larger infinity. This was an idea that clashed head-on with Cantor's own profound religious beliefs.

The continuum hypothesis

So far, we have discussed the concept of the cardinality of a set. We know it is a concept that refers to the number of elements in a set. We have also seen that it is possible to enumerate finite sets, in the sense that it is possible to assign a natural number to each element in a consecutive manner. Furthermore, when dealing with sets with an infinite number of elements, it is possible to enumerate them by means of what we have called biunivocal correspondence, a correspondence that assigns a natural number to each element of the set. We call sets where this is possible 'enumerable'. However, we have also seen that there are sets which are not enumerable and, in order to provide a way of referring to the 'quantity' of elements they contain, we have made use of the concept of cardinality. Hence, the cardinality of a set is, strictly speaking, not a number, it is a concept associated with the idea of

numeric magnitude. Herein lies an extraordinarily ingenious trick for determining the size of a set. In fact, the trick consists of comparing sets using well-defined rules that allow us to be sure when both sets are equally large and when they are not, regardless of whether they are finite or infinite.

THE FREEDOM OF MATHEMATICS

It was Cantor's desire for a free mathematics, and if anything, this is what he achieved. But what does it mean. His sense was that nobody or nothing (at least in so-called civilised countries) should stand in the way of a mathematical theory for philosophical or religious reasons. As an example, what are currently referred to as 'large cardinals' are sets of a size that is so monstrous as to dwarf Cantor's transfinite numbers. Their definition is highly complex, although the manner of constructing them bears certain similarities to the generation of the alephs based on considering the sequence of sets included one inside the other, and then taking the corresponding power sets.

Hence, Cantor used the term aleph-null to refer to the cardinality of the natural numbers, $|\mathbb{N}| = \aleph_0$, whereas for the cardinality of the real numbers, $\mathbb{R}$, he used an alternative symbol, c, which stands for the *continuum*. The reason is that the real numbers completely 'fill' the straight line (also known as the real line), and since this line is now a continuous sequence of numbers (it has no gaps) it can be referred to as continuous. Hence,

$$|\mathbb{R}| = c = 2^{\aleph_0}.$$

However, the aleph numbers constitute an ascending sequence in the sense that

$$\aleph_0 < \aleph_1 < \aleph_2 < \ldots$$

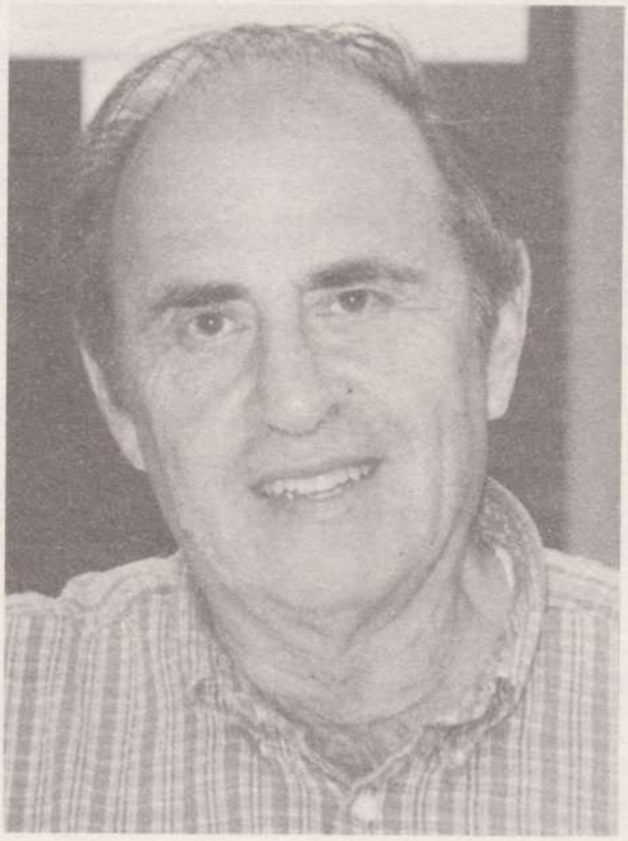
At this point Cantor asks the following question: is there a cardinality that lies between that of the natural numbers and the continuum? Somehow, he surmised that the following equality would hold.

$$2^{\aleph_0} = \aleph_1.$$

Hence, there are no sets with a 'size' that lies between that of the set of natural numbers and that of the set of real numbers. He referred to this conjecture as the *continuum hypothesis*. Cantor went to great lengths, leading him to the verge of exhaustion, to prove this result. On more than one occasion, he believed he had succeeded, but he never managed to achieve a fully satisfactory proof.

Various contemporaries of Cantor's, such as Hilbert, Russell and Zermelo, also tried unsuccessfully to prove the continuum hypothesis. The Hungarian mathematician G. Köning (1849-1913) presented a proof that the hypothesis was false at the Heidelberg Conference in 1904. Cantor always believed the proof was incorrect: he had a blind faith in his intuition, but was unable to find fault with Köning's proof. In fact, it was Zermelo who found the fault and the problem remained open. In 1900, Hilbert included it in his famous list of the 23 most important problems without a solution.

In 1963, the North American mathematician Paul J. Cohen (1934–2007) provided a proof based on Gödel's results on axiomatic consistency, that the continuum hypothesis could be true or false depending on the system of axioms chosen as the basis of set theory. This led him to a situation incredibly similar to that established by



In 1963, the American mathematician Paul J. Cohen proved that the continuum hypothesis, one of the great open problems in mathematics, cannot be proven using set theory's current axiomatic framework.

Euclid's famous fifth postulate on parallel lines – only one parallel line can be traced through a point lying outside a line. Depending on the geometry used, the postulate was true (Euclidean geometries) or false (as is the case with hyperbolic geometry).

In spite of all this, there are still those who believe the matter is not decisively settled, since a new collection of axioms to reinforce set theory could render the continuum hypothesis answerable. Furthermore, while this does not occur, we cannot be sure we have a clear idea of what a real number actually is.

Chapter 6

Cantor's Inferno

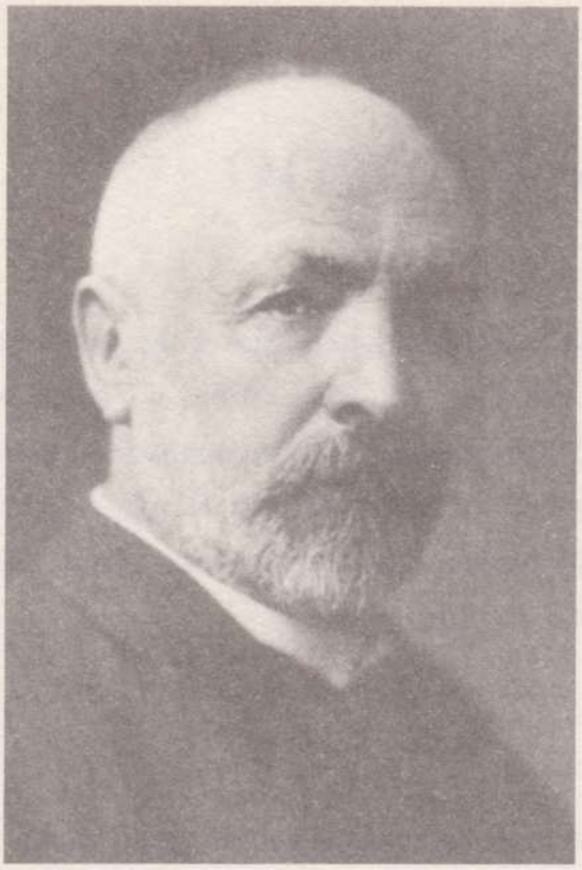
When explorers discover new lands, new places that must be charted on maps and added to geography books, their discovery comes at a price. No discovery comes for free. Some are rewarded with the glory and recognition of their achievement, whereas others slip into the obscurity of death without even having achieved the consolation of knowing their contribution was worthwhile. Georg Cantor's exploration of infinity took a heavy toll on him.

Early years

Georg Cantor was born in St Petersburg, Russia, on 3 March 1845. His father, Georg Waldemar Cantor, came from Denmark, but moved to the then Russian capital at an early age. There he set up a flourishing business importing textiles, later turning his attention to the stock market. He came to amass a considerable fortune, albeit one based on sacrifice, determination and knowledge, values he tried to inculcate in his children, whom he educated in accordance with his strict Lutheran morality. He married Maria Anna Böhm, a Russian-born Catholic, and daughter of the conductor of the St Petersburg Opera. Georg Woldemar also came from a family with a deeply rooted musical tradition, and hence it should come as no surprise that his children were also schooled in the art.

Cantor was the first-born of four brothers. In the early years of his childhood, he was educated by private tutors, until 1856, when he began primary school in St Petersburg. Cantor would always remember the first years of his childhood in Russia as the happiest of his life.

In 1856, as a result of an illness affecting his lungs, Cantor's father was forced to abandon the harsh Russian winter and move with his family to Germany. After a short stay in Wiesbaden, they finally settled in Frankfurt. Georg received his first education as a boarder in a private college located in Darmstadt, a small city close to Frankfurt, from which he graduated in 1860. By that time he had already shown an exceptional aptitude for mathematics, especially when it came to trigonometry. However, his father was unable to see a future for somebody who dedicated their



Georg Cantor, the founder of set theory, is regarded as one of the most famous mathematicians in history.

life exclusively to mathematics. He suggested that his son study engineering. As was usual, Cantor acceded to the desires of his father, and at the age of 15 entered the Tiesbaden Institute.

Cantor received many letters from his father, the majority of which aimed to boost moral strength based on religious principles. Among these, a letter he sent on 25 May 1862, stands out. In it, among other things, he writes:

> "[…] It is frequent for the most promising individuals to fail after putting up a weak resistance to the difficulties that follow from their entrance into practical affairs. Once their courage has been exhausted, they go to waste and even, in the best cases, will become nothing more than ruined geniuses... Believe me, dear son, your most sincere, genuine and experienced friend, that firm heart that must live within you is a state of mind of true religiosity... To prevent all the problems and difficulties we shall inevitably encounter as a result of envy and slander by open or secret enemies in the face of our aspirations to succeed in the activity of our speciality or business, to combat

them successfully, you must above all acquire the greatest possible quantity of knowledge and technical ability... I will close with these words: your father, in reality your parents and all the other members of the family, both in Germany, as well as Russia and Denmark, have set their sights on you, as their eldest son, and hope that you will become a brilliant star on the horizon of science. May God give you strength, health, an upright character and his greatest blessings. And may you follow his paths for ever. Amen!"

This letter contains certain prescient hints at what would become Cantor's agitated professional life. His father was doubtless an intelligent man who could sense not only his son's strong attraction to mathematics, but also his restless and creative spirit, and wished to prepare him for the challenges he might encounter. Proof of this is that very year he authorised his son to begin his studies in mathematics. Cantor sent the following letter to express his gratitude:

"Dear father, if only you could realise the great pleasure your letter has given me. It settles my future... Now I feel happy seeing that you will not be displeased if I follow what I enjoy the most. I hope that you, dear father, will live to find pleasure in my conduct, given that my soul, all my being, lives in my vocation. What a man desires, and that to which his internal compulsion pushes him, he shall comply with."

These words of gratitude, which look more like those of a young man whose family has finally accepted that he takes holy orders, expresses the profound gratitude Cantor feels for receiving his father's blessing to begin his studies in mathematics. Certain biographers agree that unconditional obedience to his father was one of the most important factors that caused Cantor to be continuously dogged by great professional insecurity in the world of academia.

In 1862, he began his studies in mathematics, philosophy and physics at the University of Zurich. However, this was a relatively short-lived period, his father's death in June 1863 resulting in him moving to the University of Berlin. It is curious that from this point onwards, Cantor made no further mention of his father.

Until the start of the 19th century, France had led the world of mathematics in terms of theory. However, around the time that Cantor moved to the University of Berlin, Germany was taking over. There he was taught by famous mathematicians

such as Kronecker, Kummer and Weierstrass. The last of these would come to exercise a great influence on Cantor, while the first, who introduced him to number theory, would become his greatest foe.

The majority of the work Cantor carried out during that period was in the fields of arithmetic and algebra. In the summer of 1866, Cantor had the fortune to visit the University of Göttingen, one of the most prestigious centres of mathematics in the world. Upon his return to Berlin, he formed part of a group of young mathematicians that met weekly in a bar to discuss mathematics in a more relaxed environment than the university classrooms.

In 1867, Cantor was awarded his doctorate for a thesis in which he carried out an in-depth study of Gauss' *Disquisitiones Arithmeticae*. The prologue provides an affirmation that prefigures what would become the restless mind of one of the most important mathematicians in history: "In mathematics, the art of proposing problems is much more stimulating than solving them." His doctoral thesis allowed him to obtain the position of *Privatdozent* at the University of Halle, an academic post where payment was provided directly by undergraduate students, such that his wages depended on the number of students attending his classes.

Halle was a small city close to Leipzig. It lacked the prestige of larger universities such as Berlin or Göttingen, something Cantor was well aware of. However despite never giving up in his attempts to leave Halle, he would come to spend the rest of his life there.

It was in 1873 when Cantor first suggested the possibility that different types of infinities could exist. He guessed that among the sets of natural and real numbers, there could be not only qualitative differences, but also quantitative ones. The former were obvious: the set of natural numbers is enumerable, whereas the set of real numbers is not. The next step, to show that the infinity of the real numbers is greater than the infinity of natural numbers would represent a milestone, not only in Cantor's own thought process, but also in the history of mathematics itself.

In 1874, a first proof appeared in the *Crelle's Journal*. We must bear in mind that at that time it was not even possible to discuss sets with the same fluency as we have done in this book. Cantor's first publication appeared in 1878 under the title *A Contribution to Set Theory*, also published in *Crelle's Journal*. In addition to containing a completely unexpected result on algebraic numbers, the paper was to usher in a new era in the history of mathematics. It already points to the ideas, still in their early stages, of transfinite cardinalities. However, instead of resulting in academic recognition that would allow Cantor to find a better position to

The University of Halle, where Cantor taught from 1872. The mathematician lived in this small German city until his death.

continue his research, it represented the beginning of a terrible ordeal in which certain mathematicians, such as his old teacher Kronecker, would make use of their academic standing to belittle Cantor professionally, a defeat that would have serious psychological repercussions.

Scientific journals

In 1826, August Leopold Crelle (1780–1855) founded *Journal für die reine und angewandte Mathematik (The Journal for Pure and Applied Mathematic)*. The title expressed a desire to strengthen the thematic unity of mathematics. It avoided referring to mathematics in the plural in a mediaeval or Renaissance legacy. It should be kept in mind that mathematical journals formed part of a wider field of publishing, that of scientific journals.

The first ever scientific journal was funded by the Royal Society of London and the dissemination and character of the early publications would remain under the umbrella of scientific societies. There are a number of interesting aspects regarding

the early publications exclusively dedicated to mathematics, such as *Gergonne's Annals* and *Crelle's Journal*. First of all, the articles were considerably shorter than those published in books. Secondly, old texts were not published. Innovation and originality were the conditions for publication. In addition joint works were published for the first time, instead of just those by a single author as had been the case until that moment in time.

The basic goal of the mathematical societies was to cover as much ground as possible and to ensure the historical permanence of the journals they published, providing them with the material they required for their dissemination. With the passing of time, it became clear that such survival was impossible without the support of official institutions. It was inevitable that all these bodies would be subject to certain social and political influence, since they became a symbol of cultural identity for the countries that supported them. This proved a double-edged sword, since while it is certain that on the one hand it represented significant backing, on the other it had the drawback that the potential internationality of science could be restricted by tight borders. Additionally, the controlling bodies that accepted publications may have lacked the desired scientific objectivity. In the long run, time has shown that

A SICILIAN MATHEMATICIAN

Giovanni Battista Guccia.

One of the first mathematical societies was, perhaps surprisingly, founded in the city of Palermo, Sicily, thanks to the publication *Rendiconti del Circolo Matematico di Palermo*. It was founded by the Italian mathematician Giovanni Battista Guccia (1855–1914), who based the authority of his publications on the fact that the society was located in one of the countries with the greatest mathematical pedigree in history. The fact is this good upbringing, together with various prizes instated by Guccia himself, led distinguished mathematicians to submit their work to the Sicilian association, which soon achieved unexpected fame, ranking near the top in the international league of mathematical societies.

the presence of mathematical societies acts as a barrier to certain innovative work that does not conform to certain canons established by a closed community, often governed by non-scientific interests. By means of example, let us note that two thirds of all articles on mathematics published in 1900 were published in non-mathematical sources.

Among the first scientific societies that began to appear in the second half of the 19th century, the most important mathematical societies were, in order of appearance: the Moscow Mathematical Society (1864); the London Mathematical Society (1865); the French Mathematical Society (1872); the Mathematical Circle of Palermo (1884); the American Mathematical Society (1888); and the German Mathematical Society (1890).

INDISCRETE MATHEMATICS

The scientific journal founded by Henry Oldenburg in 1665, *Philosophical Transactions*, has been published continuously since that date to the present day, with only two interruptions. One was caused by an epidemic of the plague in London, and the other by the illness of Oldenburg, an otherwise tireless worker. His enthusiasm was such that he came to write five letters a week, a firm believer that science should have neither barriers nor borders, leading him to continue publishing his letters even during times of war. This was regarded as a complete lack of discretion during politically sensitive times, and he was locked up for a summer in the Tower of London.

The controversy of the infinite

Kronecker remarked on one occasion that "God created the first 10 numbers; the rest is the work of man," thus affirming his vision of the task of mathematics. Everything should be constructed using these known elements, which have preferably been defined in a process involving only a finite number of steps. In other words, Kronecker did not wish to know anything about actual infinity. On one occasion he stated that infinity should be rejected "as a pernicious futility inherited from outmoded philosophies and confused theologies; he [the finitist] can get as far as he likes without it..."

Thus does Kronecker define himself as a clear follower of so-called 'finitism'. And while we are applying labels of the type '-ism', also so-called 'operationism', which represents a form of rejecting any reasoning not backed by a concrete,

well-defined operation. In this respect, he advocated monitoring by the recognised academic authorities to ensure the "wealth of its practical experience with healthy and interesting problems will give a new meaning and new impetus to mathematics. Unilateral and introspective mathematical speculation leads to sterile fields." (This last phrase was a clear reference to the work of Cantor.)

We should keep in mind that Kronecker was one of the editors of *Crelle's Journal*, and hence it should come as no surprise that in 1877 he opposed the journal publishing any work written by Cantor. This opposition led to what could be regarded as a simple scientific disagreement turning into personal insults, branding Cantor as a renegade, charlatan or corrupting influence of studious youth.

Here we should not forget that Cantor was Kronecker's most promising student and it is logical to assume that the attitude of his former teacher hurt him deeply and represented a psychological burden from which he would never break free.

Dedekind

Julius Richard Dedekind (1831–1916), who was born in Brunswick, Germany, was the fourth child from a well-off family, who dedicated the majority of his life to mathematical research. He can be regarded as an algebraist who studied the foundations of analysis, choosing sets and their applications as his building blocks.

Weierstrass, Cantor and Dedekind carried out independent research into the real numbers. The work of the last two authors represents the classical constructions that appear in textbooks today. Cantor's work, by adopting a more set-theoretical schema, was closer to Dedekind's thought, especially on account of the way both tackled the over-arching theme, with an approach grounded more in philosophy than mathematics when it came to the continuity of space. Both Cantor and Dedekind claimed there was no possibility of demonstrating such continuity. At best, we could hope to adopt it as a postulate.

In 1872, when holidaying in Switzerland, Cantor met Dedekind, one of the few mathematicians of his time (but not the only one) with whom he maintained communication through letters, based on a mutual trust and respect. Reading the correspondence between Cantor and Dedekind during the period 1874–1884, it is possible to be present at the birth of set theory. However, it is curious to note that the majority of the most important writings by the latter hardly mention the word *set*. This is due to the fact that Dedekind believed the path revealed by Cantor was already one that could be travelled and he focused on the application of the theory.

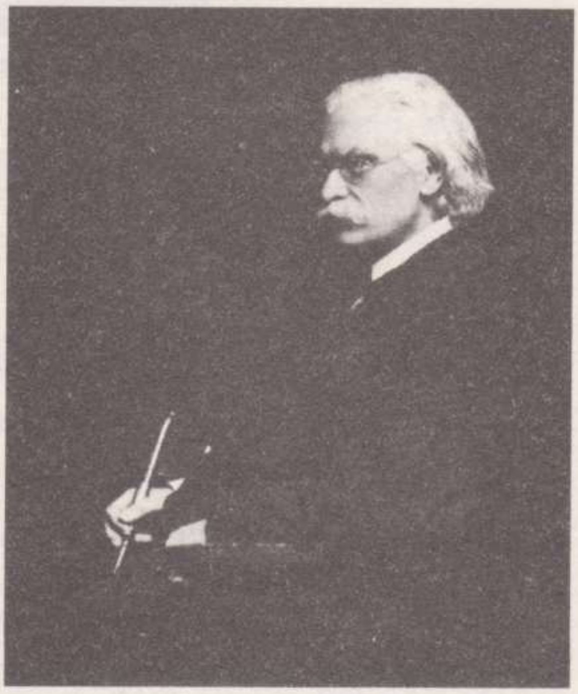
This commemorative stamp in honour of Dedekind also shows the formula that represents the decomposition of a number into prime factors.

In 1881, a vacancy arose for the chair of mathematics at the University of Halle and Cantor recommended Dedekind for the post. He did so with great gusto, writing a letter to the ministry in which he eulogised the ability of his friend. However, and in spite of Cantor's insistence, Dedekind ultimately rejected the position. The fact was that Dedekind was not driven by any form of social ambition when it came to academia. He gave classes at Collegium Carolinum for 30 years, the same institution for which his father and grandfather had worked. However, to make things worse, the ministry ended up giving the post to a candidate recommended by Kronecker. These events resulted in the cooling of the relationship between Cantor and Dedekind, and they stopped writing to each other for 16 years, breaking off relations until 1899, when Cantor finally broke the silence.

Mittag-Leffler

At the same time as the relationship between Cantor and Dedekind was about to freeze over, a new figure appeared on the horizon who would achieve a certain social standing in the scientific world and support Cantor during some of his most difficult times. This was Gösta Mittag-Leffler (1846–1927). The Swedish mathematician is better known for his attempts to disseminate the work of great mathematicians than for his own contributions. His marriage to a rich heiress allowed him to spend his time, effort and money on founding a new journal, which he did in 1882: *Acta Mathematica*, which would go on to acquire significant prestige in the international community. Cantor and Mittag-Leffler soon bonded and the latter agreed to translate the majority of the articles proposed by Cantor. A small group was set up for

translating and reviewing the texts into French, headed by Charles Hermite. As we mentioned in Chapter 5, he made notes on the proof of the transcendence of the number *e*, although the final versions were revised by Cantor himself. The publications in *Acta* would come to represent a highly significant contribution to the new theory of transfinite numbers. However an unfortunate incident that occurred due to the publication of *Principles of a Theory of Order Types* would put an end to this backing. Cantor had been struggling to provide a proof of the continuum hypothesis and had not reached an acceptable result. The aforementioned work established firm foundations for what Cantor regarded as a push to set theory that would serve to facilitate the proof. Mittag-Leffler postponed the publication of this paper for over a year, claiming that not only did it fail to provide a proof of the continuum hypothesis, but that it would also meet with fierce opposition from the scientific community. He argued that Cantor used transfinite numbers in a language that had still to be accepted by mathematicians, in addition to containing, together with the series, philosophical concepts that were foreign to mathematical reasoning. In his own words, Cantor regarded this rejection as a 'genuine catastrophe', both mathematically and personally. Furthermore, he sensed the presence of the 'black hand', the term he used to make reference to the group of mathematicians in Berlin who were reluctant to accept his theories, and which at that time included Kummer, Weierstrass and Kronecker. This was particularly the case with the last member of the group, with whom, as we have already mentioned, Cantor sustained one of the most bitter disputes in the history of mathematics.



A photograph of Gösta Mittag-Leffler taken in 1916.

Cantor the eccentric

In March 1874, on one of his frequent trips to Berlin, Cantor met Vally Guttman, a friend of his sister Sophie. They married in August that same year. Vally was a young, passionate lover of music, for which Cantor had always had a soft spot. However, aware of his weaknesses, he explained to her prior to getting married that "You will have to learn that, without apparent causes, I can be defeated by the weight of life." Regardless of this warning, their marriage was by all accounts a happy one. They had four sons and two daughters. Cantor, who had inherited enough money to not have to worry about financial pressures, decided to build a house in Halle. For a long time, he had resigned himself to staying at the small, insignificant university in that small, insignificant city and gave up trying to obtain a better academic position at the University of Berlin.

In 1885, weary of his futile attempts to prove the continuum hypothesis, and also deeply disenchanted and frustrated by the fact that the mathematical community had practically ignored him, he entered a period in his life in which his mathematical research was relegated to the background. In 1889, in his attempts to prove that the work of William Shakespeare (1564–1616) had in fact been written by Francis Bacon (1561–1626), he devoted himself to the study of the controversial English philosopher, lawyer and politician who had attempted important scientific reform. In 1898, he taught a course on the 'Life and Work of Francis Bacon' at the University of Halle and was expelled from the Shakespearean Society that same year. Cantor came to amass a significant collection of works by English authors from the 16th and 17th century, investing part of his fortune in doing so. He also spent a number of years studying philosophy and left behind some writings on the discipline. Here his interests were fundamentally focused on metaphysical matters and, above all, on those who made reference to order or were related to the concept of actual infinity.

On 16 December 1899, upon his return from Leipzig, where he had given a lecture on Francis Bacon, he found that his son Rudolf had died. He was 13 years old and had been a sickly child. Rudolph was convinced that he would be able to gain strength through studying the violin, an instrument for which he had a special gift, like most members of the Cantor family. On that occasion, Cantor made a surprising declaration in which he expressed his regret at having abandoned music for mathematics, that 'strange idea' that had caused him to abandon his true vocation.

Mental illness

Much has been written and speculated about the mental problems that dogged Cantor during the last years of his life. The difficulty of establishing a diagnosis is largely due to the lack of a clinical history from the time. All the evidence suggests that he suffered from what is currently known as bipolar disorder, a mental illness whose sufferers alternate between states of great excitement and phases of depression, without the apparent existence of an external trigger for the crisis. This is one of the reasons why claims that Cantor's madness is blamed on the savage attacks made by his colleagues (particularly Kronecker) are regarded as exaggerated.

Regardless, the last 20 years of Cantor's life passed with successive periods of confinement in psychiatric institutions, which he entered voluntarily. This did not prevent him from working and publishing his theories in the gaps of good health between the dark periods, the last of which came in May 1917. At that time, Germany was losing the war, and as it did so, a large part of its wealth and quality of life. The already precarious conditions of psychiatric institutes, such as those in Halle, were worsened by these circumstances. This last period of internment was the only one carried out against Cantor's will, and in his letters to friends and family he complained continuously about the cold, solitude and lack of food. In spite of the fact that by that time his theories had already achieved significant recognition among the scientific community, on 6 January 1918 he died alone in conditions that can only be described as miserable.

Cantor's theories on the infinite are regarded as some of the most revolutionary contributions made to mathematics in the last 25 centuries and many historians of science regard his set theory as one of the most brilliant works of human thought.

Surely, there is little point in trying to decide if Cantor's madness was congenital or caused by his circumstances. However, it is probable that, like in the majority of cases of mental illness, both were a factor. At any rate, Cantor lived the intense solitude of those who see glimmers of light in the darkness. One of his philosophical writings, published in 1883, contains a paragraph that can be read as a hymn to liberty, but that can also be interpreted as a cry, as a fierce recrimination against a society stifled by its own dogma:

A TRAGIC DEATH

In addition to the death of his son, one of the incidents that had the greatest effect on Cantor's personal life was the death of his younger brother Ludwin. In spite of sharing a strong emotional bond and their first years of schooling, the two obtained extremely different marks. Ludwin was not a good student and decided to dedicate himself to the business world when Georg entered the university. In 1863, Ludwin emigrated to the United States, although there is hardly any biographical information covering that period. All that is known for sure is that in 1870 he died in a mental health institute, which he had entered after succumbing to severe depression. Many suggest that this offers clear proof of a possible hereditary mental illness within the Cantor family.

> "Mathematics is completely free and its development, and its concepts are only restricted by the requirement that they do not contradict each other, and they fit with previously introduced concepts using precise definitions. The essence of mathematics is its freedom."

Cantor preferred to use the expression 'free mathematics', instead of the more common term 'pure mathematics'.

He died alone, but Cantor will never be forgotten. It was doubtless Hilbert who wrote his best epitaph: "No one will drive us from the paradise that Cantor created for us."

Infinity in the 21st century

Until the appearance of modern physics, infinity had only been present in the realms of theology and philosophy. In mathematics, its presence had come, to a certain extent, naturally. As Kronecker might have a said, the natural numbers were the 'givens', forming a series that extends to infinity. Geometry, thanks to the concept of the infinite line, also saw itself obliged to incorporate the diatribes between actual and potential infinity. However, it was infinitesimal analysis, *calculus*, in which it was incorporated and came into its own. In the words of Hilbert: "In a certain sense, mathematical analysis is a symphony of the infinite."

SETS AND NAZISM

The mathematical community decided to honour the work of Cantor, and to do so organised an international event to coincide with the celebration of his 70th anniversary, which was unfortunately prevented from taking place by the events of the World War I. Regardless, a group of German mathematicians met privately in his house to present a marble bust in honour of Cantor, which currently stands at the entrance to the University of Halle. During the Nazi period, the bust was removed as set theory was regarded as related to Judaism.

However, advances in physics and astronomy have incorporated infinity into the reality that surrounds us. Towards the start of the 20th century, astronomers offered us a universe that contained the Sun, with its planets and distant stars. Shortly after, this our Solar System had been immersed in a galaxy with many millions more solar systems. Space was expanded to grow large enough to hold many more thousands of galaxies. And why stop there? Who is to say that new, larger structures will not appear, requiring us to increase the size of the universe yet again? Is the universe infinite? There is still no answer to this question, and perhaps one will never be found.

Furthermore, as we increasingly enter into the details of subatomic structures, the infinitely small has also suddenly come to form part of the realm of physics. The atom, as such, is no longer the 'indivisible' it was for the Greeks. And it doesn't stop there. Nuclear physics has discovered particles within the atomic nucleus itself with dimensions of less than 10^{-15} metres. Up to this point, it was possible to talk of unimaginably small magnitudes, but not of infinitely small ones. However, one of the theories of physics that has offered the greatest resistance to experimental proofs – quantum electrodynamics – deals with elementary particles, such as electrons and quarks. Mathematically speaking, they are regarded as points and fit the range of points of a straight line, or the real numbers.

Perhaps one day, science will come to show that in nature there is not and has never been, a clear difference between potential and actual infinity, and that, at the heart of the matter, it is nothing more than a battle of the mind.

Appendix

The Irrationality of $\sqrt{2}$

The first record of the irrationality of the square root two dates back to a pre-Socratic philosopher from the Pythagorean school called Hippasus of Metapontum (born c. 500 BC). In addition to showing off his mathematical talent in the process, Hippasus showed great courage by tackling a subject that remained taboo in his cultural milieu. Remember as legend has it the punishment for Pythagoreans who even dared mention the existence of irrational numbers was death.

Like the majority of such proofs, including those found in certain apocryphal editions of Euclid's *Elements*, Hippasus uses the method of *reductio ad absurdum*. In modern language, the argument is as follows:

If $\sqrt{2}$ is a rational number, this means it can be expressed as the quotient of two integers of the form

$$\sqrt{2} = \frac{p}{q},$$

where this is an irreducible fraction (i.e. its numerator and denominator do not contain common factors). Squaring this gives:

$$2 = \frac{p^2}{q^2}$$

and hence,

$$p^2 = 2q^2.$$

This means that p^2 is an even number and hence so is p. This allows us to express p as a multiple of two, hence, $p = 2n$, meaning that

$$2q^2 = p^2 = (2n)^2 = 4n^2,$$

and simplifying,

$$q^2 = 2n^2.$$

Thus, q^2 is an even number and hence so is q. We have reached the conclusion that both p and q are even numbers, or rather the fraction p/q contains common factors, thus contradicting the hypothesis with which we began. This means that $\sqrt{2}$ cannot be equal to the quotient of two integers.

The first approximations of $\sqrt{2}$ contained as few as four or five decimal places. A good approximation consisting of 65 digits is: $\sqrt{2}$ = 1.41421356237309504880168872420969807856967187537694807317667973799.

Using modern computing, it is possible to obtain approximations correct to millions of decimal places.

Sets of numbers

The definition of the different sets of numbers is complicated and requires mathematical knowledge beyond the scope of this book. However, there is another way of understanding them, less rigorous but more intuitive, based on their practical applications for solving equations. The starting point is the so-called natural numbers. The natural numbers, 1, 2, 3, … are represented using the symbol $\mathbb{N}$. This is a set expressed as follows:

$$\mathbb{N} = \{0, 1, 2, 3, 4, 5, 6, 7, \ldots\}$$

Certain authors do not include zero as a natural number, something which is perfectly justifiable when we remember that the number is the result of a long and profound mathematical revolution and hence there is very little that is 'natural' about it.

Using the set of natural numbers it is possible to solve equations such as:

$$x - 2 = 0.$$

But not of the type $x + 2 = 0$, since the set does not include negative numbers. When we add the negative numbers and zero, this gives the set of integers, represented using the symbol $\mathbb{Z}$.

The remaining sets of numbers are introduced along the same lines. For example, to solve an equation of the type

$$2x + 3 = 0,$$

whose solution is $x = -3/2$ it is necessary to introduce rational numbers $\mathbb{Q}$.

For an equation such as

$$x^2 - 2 = 0,$$

it is necessary to add irrational numbers. Joining these to the rational numbers gives the set of real numbers $\mathbb{R}$.

Finally, the equation

$$x^2 + 2 = 0$$

does not have a real solution, since no real number can be the square root of a negative number. The next and final step that allows us to solve this kind of equa-

tion is the introduction of imaginary or complex numbers, whose set is represented using the symbol $\mathbb{C}$. We say that this is the final set because it has been shown that any equation with complex coefficients always has a solution (thanks to the fundamental theorem of algebra).

Each of the sets we have established has been included in the following (an algebraic extension):

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}.$$

Bibliography

Boyer, C.B., *A History of Mathematics,* Hoboken, John Wiley, 2011.

Dauben, J., *Georg Cantor,* Princeton, Princeton University Press, 1990.

Dedekind, R., *What are Numbers and What Should They Be?* Orono, Research Institute for Mathematics, 1995.

Guthrie, W.K.C., *A History of Greek Philosophy,* Cambridge, Cambridge University Press, 1990.

Kline, M., *Mathematical Thought from Ancient to Modern Times,* New York, Oxford University Press, 1990.

Mankiewicz, R., *The Story of Mathematics*, Princeton, Princeton University Press, 2004.

Stewart I., *Taming the Infinite: the Story of Mathematics from the First Numbers to Chaos Theory,* London, Quercus, 2011.

Index

Into Infinity

Discovering a world without end

If the search for knowledge is compared to a treasure hunt, a complete understanding of the theory of infinity would without doubt be the greatest prize of all. Carving a path between dogma and paradox, the concept of infinity has always been a nebulous yet tangible presence in the landscapes of ancient philosophy, medieval religion and modern science. This book follows the elusive trail of infinity through the minds of philosophers, theologians, physicists and, above all, mathematicians.